At the time of issue this report has not been published. Any author desiring to quote this paper should refer to it as "A paper of the Meteorological Research Committee (London).
M.R.P. No. 953 by Mr. D.H. Johnson.
A copy is available in the Library of the Meteorological Office."

M.R.P. 953
S.C.II/202
31st December, 1955

AIR MINISTRY
METEOROLOGICAL RESEARCH COMMITTEE

The success achieved in forecasting upper winds by orthodox and statistical techniques

by
D.H. Johnson (Meteorological Office)

1. Introduction

C. S. Durst has explored the possibility of forecasting upper winds by linear regression equations, which exploit judiciously the temporal serial correlation of the wind (1). Forecasting trials were held, during the first four months of 1955, to provide data for a strict comparison between the success of the statistical method and the successes of the techniques currently in use at the Central Forecasting Office and at London Airport. The results of the trials are analysed in this paper.

2. Techniques

Five types of forecast were made: London Airport (L.A.P.), Central Forecasting Office (C.F.O.), Prontour, Statistical and Persistence. The techniques by which these forecasts were made are explained below. Winds forecast were:

(i) point winds at Liverpool and O.W.S. Station 'J' for the 700, 500, 300, 200 and 100 mb. levels (London Airport made no forecasts for 100 mb.);

(ii) mean tailwind components (as distinct from equivalent tailwinds) over approximately great circle routes from Shannon to Gander at 500 mb., and from O.W.S. Station 'A' to London at 200 mb.

C.F.O. and L.A.P. forecasts were made before 0300 G.M.T. on each day, all forecasts being valid for the following 1500 G.M.T.

(a) Orthodox techniques

At C.F.O. and London Airport established practice in forecasting upper winds rests largely on the use of prontours (forecast contour charts), but there are differences between the techniques in use at the two stations, which it is useful to bear in mind when assessing the results of the trial. Up to 200 mb., the C.F.O. contour charts are built up from the 1000 mb. charts by the successive graphical addition of the isopleths of thickness of the layers, 1000-700 mb., 700-500 mb., 500-300 mb. and 300-200 mb. London Airport follows a similar procedure for charts up to 300 mb., but the L.A.P. 200 mb. charts are drawn directly from the 200 mb. observations without reference to the 300-200 mb. thickness pattern. London Airport makes some use of reports from civil aircraft operating at levels up to 500 mb. These reports are not studied as a routine at C.F.O.

C.F.O. prontours are produced by graphical addition of forecast thickness patterns, up to 200 mb. L.A.P. prontours are based upon forecast thickness patterns up to 500 mb.: above that level the technique largely consists of direct extrapolation of contour patterns.

/Routine

Routine prontours prepared at London Airport for operational purposes are composite charts, that is, the time of validity of the prontour varies from east to west across the chart. The trial route forecasts were of the instantaneous mean winds over the routes. Consequently, the L.A.P. forecasters elected on most occasions, to construct special fixed-time prontours for the purposes of the trial.

Prontours are constructed to meet a schedule which allows a certain balance to be attained between accuracy and speed of production. At C.F.O. uniform attention is given to all parts of the area to which the prontours relate. In framing a specific forecast, the C.F.O. forecaster may have been led to amend the geostrophic value which the prontours imply for one of the following reasons:

(i) to allow for the 3-hr. difference between the time of validity of the routine prontour (1200 G.M.T.) and the time of reference of the trial forecasts (1500 G.M.T.);

(ii) to allow for ageostrophic effects;

(iii) as a result of more detailed consideration of conditions near to the reference point of the forecast;

(iv) in order to reduce the risk of making very large errors by avoiding the prediction of large changes or unusual values.

The value written down by the practising forecaster has been taken as the "C.F.O." forecast. A geostrophic wind measured from a C.F.O. prontour has been termed a "prontour" forecast. Prontour forecasts were obtained after the event, in the Forecasting Research Division.

(b) Statistical forecasts

The statistical forecasts were prepared in the Forecasting Research Division. They were based upon the latest available wind observation, i.e., upon the 2100 G.M.T. observation, or in the absence of a 2100 G.M.T. observation, upon a 1500 G.M.T. wind or a geostrophic estimate of the 1500 G.M.T. wind. So that no use should be made of observations which were not available to the orthodox forecaster, a log was kept of the times of receipt, in the forecast room at C.F.O., of relevant observations: those received later than 0300 G.M.T. were discarded.

The basic principles of the statistical method are discussed in Geophysical Memoir No. 93. Details of how the statistical forecasts were made are given in Appendix I to this report. Values of the regression co-efficient etc., were provided by C. S. Durst.

(c) Persistence forecasts

The forecasting success of a current wind observation provides a useful standard against which to judge the value of more subtle techniques. For both route and point winds estimates of wind at 1500 G.M.T. or 1500 G.M.T. wind observations have been used as forecasts valid for the following 1500 G.M.T. These are referred to as persistence forecasts.

3. Winds over the routes

(a) Estimated "actuals"

Estimates of the mean wind over the route, the "actuals", against which the route forecasts were verified, were made by three methods:

(i) London Airport used a template to obtain from their working charts the average geostrophic wind over a band, approximately 300 miles wide, centred on the route;

(ii) the F.R. Division averaged spot winds, estimated at a series of points equally spaced along the route. Both geostrophic and observed winds were taken into account;

(iii) statistical estimates were made using equation (B) of Appendix I. Means and standard deviations for each type of estimate are given in Table 1.

Table 2 contains details of comparisons between the estimates. The agreement between the C.F.O. and L.A.P. estimates was evidently close: the mean differences were negligible, the correlation co-efficients were high, and the mean square differences between the estimates were small in comparison with the variances of the estimates themselves. The C.F.O. and L.A.P. estimates were averaged to give a further value, which has been termed the "best" estimate. This latter value was the one used in verifying the forecasts. The statistical estimates agree well with the other three, although they contain a systematic error of about 2 kt. for the route Shannon-Gander.

(b) Properties of the statistical forecasts

In assessing the value of the statistical forecasts, and in discussing their behaviour, it is necessary to refer to a number of formulae, most of which are not quoted in published work. These are listed below. The derivation of one of the relationships is given: the remainder are readily derived by similar algebra.

The basic prediction equation is of the form

$$F_t - U = P(V_o - U) \qquad \ldots (1)$$

where V_o = wind observed at time t = o, F_t = wind forecast for time t, and U and P are constants. Durst (1) shows that for such a prediction equation the forecast errors will be a minimum if U is identified with $\bar{V}$, the mean wind for the period over which forecasts are made, and P is identified with R_t, the correlation between observations made t hours apart. Equation (1) and the following relations apply both to route wind forecasts, which are scalar quantities, and to the vector point winds. For vector winds the correlation co-efficients refer to stretch vector correlations. Values for U and P have to be chosen before the event and so they will not generally be identical with $\bar{V}$ and R_t.

From equation (1), taking mean values over n occasions,

$$\bar{F}_t = U + P(\bar{V} - U) \qquad \ldots (2)$$

and

$$\bar{F}_t - \bar{V} = (U - \bar{V})(1 - P) \qquad \ldots (3)$$

The variance of n statistical forecasts may be written as $\frac{1}{n}\sum(F_t - \bar{F}_t)^2$, which on substitution for F_t and $\bar{F}_t$ from (1) and (2) becomes, $\frac{1}{n}P^2\sum(V_o - \bar{V})^2$. Thus, if σ_F is the standard deviation of the statistical forecasts, and σ is the standard deviation of winds observed during the period of the trial, we have,

$$\sigma_F = P\sigma \qquad \ldots (4)$$

P is less than one, so the standard deviation of the statistical forecasts should be appreciably less than that of the observed winds.

The following relationships may be similarly obtained.

/Root

Root-mean-square error of statistical forecasts =

$$\left\{\sigma^2\left(1-2PR_t+P^2\right)+\left(U-\bar{V}\right)^2(1-P)^2\right\}^{\frac{1}{2}} \qquad \ldots (5)$$

Standard deviation of statistical forecast errors =

$$\sigma\left(1-2PR_t+P^2\right)^{\frac{1}{2}} \qquad \ldots (6)$$

Root-mean-square error of persistence forecasts =

$$\left\{2\sigma^2\left(1-R_t\right)\right\}^{\frac{1}{2}} \qquad \ldots (7)$$

Co-efficient of correlation between a t-hr. statistical forecast and the wind at time t =

$$\frac{1}{n}\sum\left(F_t-\bar{F}_t\right)\left(V_t-\bar{V}_t\right)\Big/P\sigma^2 = R_t \qquad \ldots (8)$$

Serial correlation in statistical forecast errors =

$$\frac{\sigma^2\left(P^2R_t-P+R_t-PR_{2t}\right)+\left(U-\bar{V}\right)^2(1-P)^2}{\sigma^2\left(1+P^2-2PR_t\right)+\left(U-\bar{V}\right)^2(1-P)^2} \qquad \ldots (9)$$

Serial correlation in persistence forecast errors =

$$\left(2R_t-R_{2t}-1\right)\Big/2\left(1-R_t\right) \qquad \ldots (10)$$

(c) Route forecasts

In forecasting winds over the routes, both C.F.O. and L.A.P. used templates to obtain from the prontours a mean geostrophic wind. The "prontour" forecasts were the average of spot geostrophic measurements made at points along the route. The F.R. Division's estimates of mean winds over the routes were used as persistence forecasts.

Table 3 gives the mean values and standard deviations of the five types of forecast, together with statistics relating to the forecast errors. An important statistic for assessing the success of wind forecasts is the root-mean-square error. Values of the r.m.s. error are set out in Table 3, but for this analysis it was convenient to break the r.m.s. error down into its component parts, (i) bias and (ii) standard deviation.

(i) Bias. The term bias has been used to denote the systematic error (i.e., the difference between the mean forecast value and the mean actual). Over the route Shannon-Gander the bias in the orthodox forecasts was negligible, but in the statistical forecasts there was a bias of nearly 6 kt. Equation (3) shows that the bias, $(\bar{F}_t - \bar{V})$, in the statistical forecasts is proportional to the difference $U-\bar{V}$. Table 4 contains, for comparison, values of U and $\bar{V}$: $U-\bar{V}$ averaged 17 kt. during the four months of the trial. For the route O.W.S. 'A' - London the statistical forecast bias of 2 kt. can be explained similarly. The 2 kt. biases in the orthodox forecasts of wind over the 200 mb. route were tested

/for

for statistical significance.* It was found that the C.F.O. bias was significant at the 3% level and the L.A.P. bias was significant at the 6% level, which suggests that there is a real tendency for the orthodox forecasters slightly to under-estimate the 200 mb. tailwinds.

It is clear from Table 3 that in every case except that of the 500 mb. statistical forecasts, the bias made a negligible contribution to the r.m.s. error.

(ii) Standard deviations of the forecast errors. The following comparisons can be drawn between the standard deviations of forecast errors given in Table 3.

(i) The difference between the L.A.P. and C.F.O. values for 500 mb. was negligible but at 200 mb. the L.A.P. value exceeded the C.F.O. value by 1.3 kt. The 200 mb. difference was found to be significant at the 7% level.

(ii) The C.F.O. and statistical values differed by 1.4 kt. at 500 mb. (11% significance level), and by 1.9 kt. at 200 mb. (2% significance level), the C.F.O. value being the smaller in each case.

(iii) Correlation co-efficients. The last column of Table 3 gives the correlations of the forecasts with the actuals. The orthodox forecasts obtained higher correlation co-efficients than did persistence or statistical forecasts.

(iv) Frequency distributions of error. Figure 1 contains histograms which give the frequency of occurrence of errors of given magnitudes. The histograms show clearly the effects of bias. Table 5 shows the frequencies of occurrence of errors greater than specific magnitudes. The figures given in the last two columns of Table 5 indicate that there was little to choose between the orthodox and statistical techniques as far as the avoidance of extreme errors was concerned, although conclusions drawn from the last two columns alone could not be regarded as having a sound statistical basis. In the remainder of the table, frequencies for statistical forecasts exceed those for the orthodox forecasts.

(v) Individual errors. The graphs in Figures 2 and 3 give daily values of C.F.O., L.A.P., and statistical forecasts for comparison with the estimated actuals.

The L.A.P. and C.F.O. graphs suggest that, while there was a high correlation between the two forecasts, the forecast errors were not correlated to the same degree. On calculation, the co-efficient of correlation between the C.F.O. and L.A.P. errors was found to be .61 for the 500 mb. route and .58 for the 200 mb. route. It is thus possible to reduce the standard deviation of the forecast errors by combining the two forecasts, the theoretical reduction amounting to about 10 per cent.†

The statistical forecast graph shows the systematic errors to which the regression equation gave rise over short periods. See, for example, values for the first half of January at 500 mb. Table 6 gives details of occasions when errors greater than 27 kt. were made by one or more of the forecasts. On nine of the fifteen occasions when persistence gave a large error, either C.F.O. or L.A.P. committed an error greater than 20 kt.

/4.

* The tests of significance, which were complicated by serial correlations in the forecast errors, are discussed in detail in Appendix II.

† It is also theoretically possible to improve the statistical forecasts by basing the regression equation on the orthodox forecasts, instead of on the latest actuals.

4. Winds at a point

Point forecasts were normally verified against wind observations. In the absence of a wind observation an estimate was made from the appropriate C.F.O. working chart. At 500 mb., one wind was estimated for Liverpool and eleven for O.W.S.'J'. At 100 mb. 14 winds were estimated for Liverpool and 77 for O.W.S.'J'.

Mean values of the south and west wind components for each type of forecast, corresponding standard vector deviations and statistics relating to the forecast errors are listed in Table 7.

The C.F.O. forecast winds tended to have smaller standard vector deviations than did the observed winds, especially at 300 mb. This did not apply to C.F.O. route forecasts.

(a) Bias. For reasons similar to those discussed under section 3(c) for route forecasts, the statistical point forecasts contained a small bias. The bias in the orthodox forecasts was negligible and for all types of forecast the contribution of the bias to the r.m.s. vector error was small.

(b) Standard vector deviation of errors. The following features emerge from a comparison of the standard deviations of forecasts errors, given in Table 7.

(i) There was little to choose between C.F.O. and London Airport at 700 mb. and 300 mb. At 500 mb., the L.A.P. values for both stations were about a knot lower than those obtained by C.F.O. At 200 mb. C.F.O. bettered the L.A.P. values by about a knot at Liverpool and four knots at the ship.

(ii) C.F.O. achieved smaller standard deviations of error than the statistical method for all levels up to 200 mb. At 100 mb., there was nothing to choose between C.F.O. and statistical values for Liverpool, but C.F.O. had an advantage of one knot at O.W.S.'J'.

(iii) C.F.O. failed to improve upon their prontours up to 300 mb., although at 200 mb. and at 100 mb. (Liverpool) the prontours were bettered. Persistence plays a very large part in the orthodox technique of forecasting at 100 mb. and the gain over the prontours at Liverpool was due partly to the weight being given to the wind observations rather than to measured geostrophic winds.

(iv) In assessing the improvement upon persistence, it must be borne in mind that the values given in Table 7 refer to 24-hr. persistence forecasts. The forecaster could make use of 2100 G.M.T. observations from Liverpool on 118 out of 120 occasions at 200 mb., and 110 of these observations were available at 100 mb. (Corresponding figures for the ship are: 200 mb., 59 occasions; 100 mb., 25 occasions). Making use of equation 8, which shows that the serial correlation in the statistical forecasts should be identical with the serial correlation in the observations on which they are based, it is possible to estimate that if 2100 G.M.T. observations had been used for the persistence forecasts whenever they were available, the persistence errors would have been reduced by about ten per cent. At 100 mb., persistence would have attained approximately the same success as the orthodox and statistical methods, although at all other levels the forecasters would still have shown substantial gains.

Table 8 gives levels of statistical significance of the differences between pairs of the standard deviations of error.

(v) Frequency distributions of the errors. Figure 4 contains histograms showing frequencies of vector errors in 10 kt. ranges for each forecasting technique. In Table 9 are tabulated the frequencies of occurrence of errors greater than 50 kt. and 70 kt. respectively. Neglecting persistence, it does not appear that any one method was less liable to give large errors than the others.

/(vi)

(vi) Individual large errors. In Table 10 the outstanding errors are catalogued. It is remarkable that only one of the cases listed in Table 10 is also quoted in the corresponding table for the routes (Table 6).

5. General assessment

In drawing conclusions it is necessary to assess how far the results of the trials are likely to be of general validity. The statistical significance tests were made with this end in view, but we must also consider whether the conditions obtaining during the trial were such as to affect the representativeness of the samples of forecast errors. In Table 4 the mean winds, the standard vector deviations and serial correlations for the period of the trial are compared with longer period values and with the corresponding statistical prediction parameters. The column headed "Frequency of 1955 departure" gives the percentage of years when a departure of the mean wind from normal, equal to or greater than the departure which was observed in 1955, can be expected to occur. For reasons given at the end of Appendix II, these percentage frequencies must not be taken too literally, but they do give grounds for the belief that while the mean winds over the British Isles, during the period of the trial, were not particularly abnormal, the flow over parts of the Atlantic was far from normal. For O.W.S. 'J', then, and for the 500 mb. route, the biases in the statistical forecasts must be abnormally large. For the point winds, the effect of this was practically negligible in comparison with the non-systematic forecast errors, but for the 500 mb. route, the bias in the statistical forecast was relatively so large as to cancel out the expected gain over persistence. In view of the abnormality of the trial period, the poor success of the statistical forecasts for the 500 mb. route must be considered to be unrepresentative. While the value of U, used for the statistical forecasts, cannot be expected to agree precisely with long period means, it is considered unlikely that, if statistical forecasts were made over a long period, there would be a sufficient difference between U and $\overline{V}$ to give rise to appreciable bias. (Nevertheless, if the use of the statistical method in making special forecasts over a very limited period were contemplated, then the probable bias might require assessment). For the above reasons it is considered that a comparison of the standard deviations of the forecast errors gives a better guide to the relative success of the different techniques than a comparison of the r.m.s. errors, for the 500 mb. route.

Comparison of values given in the last two columns of Table 4 shows that whilst, for the point forecasts, P_{24} was taken as .55 at all levels, the observed serial correlations increased with height. It is considered that the trial period was not abnormal in this respect and that above 300 mb., P_{24} could, with advantage, be increased. Long period values of the serial correlations are not known, but even if P had been given values equal to the serial correlations actually observed, the statistical forecast errors would not have been reduced by more than half per cent at 300 mb. and five per cent at 100 mb.

At O.W.S. 'J' the 1955 standard deviations are somewhat less than the normals given by Geophysical Memoir No. 85. Results given for the ship are less accurate than those obtained for Liverpool, partly because the ship observations themselves were sometimes of dubious accuracy and partly because, at the higher levels, a large percentage of the "observed" winds were, in fact, estimates made from the working charts. It is known that both orthodox and statistical forecasts achieve slightly smaller successes relative to persistence during periods when the standard deviation of the observed winds is below normal, but there is no reason to suppose that either technique gains in relation to the other.

It can be concluded, then, that a comparison of the standard deviations of the forecast errors for the trial period is likely to give a fair indication of the relative success, over longer periods, of the different techniques. The results given in Table 8 regarding the comparison between the standard

/deviations

deviations can therefore be summed up in the following terms.

(i) the C.F.O. errors were smaller than the statistical forecast errors over the route and at Liverpool for levels up to 200 mb., the significance levels of the differences varying from 2% to 11%. Taken together these results provide sound evidence that the orthodox method of forecasting as applied in everyday practice is capable of giving significantly smaller forecasting errors than the statistical method for levels up to 200 mb. At 100 mb. there is nothing to choose between the two techniques.

(ii) there is no significant difference between accuracies attained at London Airport and C.F.O. for levels up to 300 mb. At 200 mb., where C.F.O. and L.A.P. techniques differ, there is a reasonably strong suggestion that the C.F.O. technique is superior to that employed by London Airport.

6. Monthly values

Monthly values of the statistics relating to the forecast errors have not been discussed in this paper, but they are held for reference in M.O.21. For the routes, data relating to the success of the forecasts when compared with the individual C.F.O. and L.A.P. estimates of the actuals, are also available in M.O.21.

7. Summary of conclusions

(1) Orthodox techniques of forecasting are capable of giving, under operational conditions, smaller forecasting errors than the statistical technique proposed by C. S. Durst at levels up to 200 mb. There is no evidence that one technique is superior to the other in avoiding extreme errors.

(2) Orthodox and statistical techniques both give appreciably smaller errors than a persistence forecast at levels up to 200 mb. At 100 mb. neither technique improves significantly upon a persistence forecast over the British Isles, but both techniques effect some improvement at Ocean Weather Ship Station 'J'.

(3) C.F.O. and London Airport attain similar degrees of success in forecasting winds up to 300 mb. There is some suggestion that the (modified) thickness technique employed by C.F.O. at 200 mb. is superior to the technique of direct extrapolation practised by London Airport at that level.

(4) The C.F.O. forecasters do not better geostrophic measurements made from their prontours, except for forecasts of wind at a point at 200 mb. and 100 mb.

8. Acknowledgements

The statistics given in Tables 1, 2, 3, 7 and 8 and in Appendix III were obtained by electronic computation on Messrs. Elliot Bros.' 401 machine, at the Rothamsted Experimental Station, Harpenden. Thanks are due to Dr. F. Yates, Head of the Department of Statistics at Rothamsted, for permission to use the facilities of his department, and to his staff, for their valuable assistance. We are particularly grateful to Mr. M. J. R. Healy, of Rothamsted, who designed the programmes and supervised operations.

/References

References

1.	Durst, C.S.	1954	Geophysical Memoir No. 93
2.	Lewis, R.P.W. and McIntosh, D.H.	1952	Met. Mag., Vol. 81, p.242
3.	-do-	1953	Ibid, Vol. 52, p. 239
4.	Brooks, C.E.P. and Carruthers, N.	1953	Handbook of Statistics in Meteorology, p. 323
5.	-do-		Ibid, p. 194
6.	-do-		Ibid, p. 233
7.	Johnson, D. H.	1955	Quart J.R. Met.Soc., Vol. 81, p.1
8.	Brooks, C.E.P. and Carruthers, N.	1953	Handbook of Statistics in Meteorology, p. 200
9.	Yule, G.U. and Kendall, M.G.	1937	Introduction to the Theory of Statistics, p. 403
10.	Jenkinson, A.F.	1955	Met. Mag., Vol. 84, p. 140

Tables 1 and 2

Comparison between estimates of winds over the routes for the four-month period January to April, 1955

Table 1

Route	Estimate	Mean (kt.)	Standard deviation (kt.)
Shannon	L.A.P.	-13.2	18.7
-Gander	C.F.O.	-13.5	19.2
500 mb.	"Best"	-13.4	18.9
	Statistical	-15.1	17.5
O.W.S. 'A'	L.A.P.	20.3	20.4
-London	C.F.O.	19.9	20.9
200 mb.	"Best"	20.1	20.5
	Statistical	19.7	20.1

Table 2

Route	Estimates compared	Mean difference (kt.)	Standard deviation of difference (kt.)	r.m.s. difference (kt.)	Correlation co-efficient
Shannon	C.F.O. - L.A.P.	0.3	5.9	5.9	.95
-Gander	C.F.O.-Statistical	1.6	7.6	7.7	.92
500 mb.	L.A.P.-Statistical	1.9	7.4	7.6	.92
	"Best"-Statistical	1.7	6.9	7.1	.93
O.W.S.'A'.	C.F.O. - L.A.P.	0.4	6.1	6.1	.96
-London	C.F.O.-Statistical	0.2	4.0	4.0	.98
200 mb.	L.A.P.-Statistical	0.6	6.6	6.6	.95
	"Best"-Statistical	0.4	4.6	4.6	.97

Table 3 Verification of forecasts of wind over the routes. January to April, 1955.

Route	Statistic	Mean (kt.)	Standard deviation (kt.)	Bias (kt.)	Standard deviation of forecast errors (kt.)	r.m.s. error (kt.)	Correlation co-efficient
Shannon	Actual	-13.4	18.9				
-Gander	L.A.P. f/c	-13.4	18.0	0.0	11.1	11.0	0.82
500 mb.	C.F.O. f/c	-13.7	19.5	0.3	10.9	10.9	0.84
	Prontour f/c	-13.1	19.1	-0.3	11.4	11.4	0.82
	Statistical f/c	-19.2	12.3	5.8	12.3	13.6	0.77
	Persistence f/c	-13.5	19.2	0.1	13.4	13.4	0.75
O.W.S.'A'	Actual	20.1	20.5				
-London	L.A.P. f/c	18.2	19.9	1.9	12.2	12.3	0.82
200 mb.	C.F.O. f/c	18.1	19.9	2.0	10.9	11.1	0.85
	Prontour f/c	17.7	18.9	2.4	10.6	10.8	0.86
	Statistical f/c	22.3	16.1	-2.2	12.8	12.9	0.78
	Persistence f/c	19.9	20.9	0.2	14.7	14.7	0.75

Table 4. Comparison between winds observed January - April, 1955 and normal values for January - April

Location	Mean observed wind, $\bar{V}$		Statistical forecast parameter, U	Frequency of 1955 departure	s.v.d. of observed winds (kt.)		Serial correlations	
	1955 (kt.)	1948-55 (kt.)			1955	G.M. 85*	Observed 1955	Statistical value, P
Route Shannon							(24-hr.)	(18-hr.)
- Gander 500 mb.	-13	-	-30	-	18.9	-	.75	.75
Route O.W.S. 'A'							(24-hr.)	(18-hr.)
- London 200 mb.	20	-	29	-	20.5	-	.75	.80
Liverpool							(24-hr.)	(24-hr.)
700 mb.	300°8	280°14	270°13	16%	27.0	29.0	.59	.55
500 mb.	295°16	287°20	280°26	42%	39.9	37.5	.61	.55
300 mb.	295°25	294°27	285°37	> 75%	56.1	53.0	.63	.55
200 mb.	295°24	296°26	285°34	> 75%	39.1	41.0	.70	.55
100 mb.	300°20	295°19	275°25	> 75%	23.5	23.0	.79	.55
O.W.S. 'J'							(24-hr.)	(24-hr.)
700 mb.	235°9	254°21	250°18	< 1%	27.7	30.7	.58	.55
500 mb.	245°13	258°29	270°30	< 1%	35.5	39.1	.57	.55
300 mb.	255°23	263°38	270°40	7%	49.2	54.5	.56	.55
200 mb.	250°23	266°34	260°41	9%	39.9	44.0	.63	.55
100 mb.	255°17	269°23	270°27	15%	22.2	26.0	.69	.55

* Averages for the six months December - May, derived from Geophysical Memoir No. 85, as amended by M.O.9

Table 5. Frequencies of occurrence of errors greater than specified magnitudes

Route	Type of forecast	Error size (kt.)							
		2½	7½	12½	17½	22½	27½	32½	37½
Shannon	L.A.P.	93	62	32	13	6	1	0	0
-Gander	C.F.O.	98	56	33	13	6	1	0	0
500 mb.	Prontour	94	64	36	17	4	1	0	0
	Statistical	102	68	47	26	12	4	1	1
	Persistence	103	72	40	21	12	3	2	1
O.W.S. 'A'	L.A.P.	101	63	37	20	6	3	3	0
-London	C.F.O.	98	59	32	14	4	2	1	1
200 mb.	Prontour	94	52	25	13	6	2	2	1
	Statistical	103	66	43	21	8	4	1	0
	Persistence	110	77	44	30	17	8	3	0

Table 6. Occasions when large errors were made

Route	Date	Estimated actual	Forecast wind (kt.)					Forecast errors (kt.)				
			L.A.P.	C.F.O.	Prontour	Statistical	Persistence	L.A.P.	C.F.O.	Prontour	Statistical	Persistence
Shannon	Jan. 27	-5	-36	-30	-33	-43	-40	31	25	28	38	35
- Gander	Mar. 26	3	-13	-10	-12	-27	-29	16	13	15	30	32
500 mb.	Mar. 29	-6	5	22	19	- 4	- 5	-11	-28	-25	- 2	- 1
	Apr. 3	-2	-25	- 8	- 5	-33	-29	23	6	3	31	27
	Apr. 13	2	-14	-21	-21	-24	-36	18	23	23	26	38
	Apr. 18	18	2	15	20	-10	- 9	16	3	- 2	28	27
O.W.S. 'A'	Jan. 11	74	38	50	60	44	44	36	24	14	30	30
-London	Jan. 22	53	38	45	45	31	25	15	8	8	22	28
200 mb.	Feb. 16	49	34	36	31	23	14	15	13	18	26	35
	Feb. 27	33	-4	-5	-5	15	8	37	38	38	18	25
	Mar. 3	46	19	25	23	18	18	27	21	23	28	28
	Mar. 6	-19	-9	-10	-13	17	15	-10	-9	-6	-36	-34
	Mar. 8	15	-18	0	-3	-1	-16	33	15	18	16	31
	Apr. 8	3	3	0	2	25	35	0	3	1	-22	-32
	Apr. 9	38	33	30	28	9	4	5	8	10	29	34
	Apr. 10	60	40	30	26	56	40	20	30	34	4	20

Table 7. Verification of forecasts of wind at a point. January - April, 1955

(a) Liverpool

Level (mb.)	Statistic	Mean component (kt.)		Standard vector deviation (kt.)	Bias in component (kt.)		Standard vector deviation of errors (kt.)	r.m.s. error (kt.)	Correlation co-efficient
		S.	W.		S.	W.			
700	Actual	-3.8	7.0	27.0					
	L.A.P.	-3.1	8.2	28.2	-0.7	-1.2	19.0	18.9	0.77
	C.F.O.	-2.9	8.3	26.5	-0.9	-1.3	18.5	18.5	0.76
	Prontour	-2.1	7.3	26.9	-1.7	-0.3	18.8	18.8	0.76
	Statistical	-2.6	8.7	17.2	-1.2	-1.7	20.1	20.1	0.67
	Persistence	-3.8	6.8	27.0	0	0.2	24.5	24.4	0.59
500	Actual	-7.4	15.3	39.9					
	L.A.P.	-6.7	15.2	39.3	-0.7	0.1	25.5	25.4	0.79
	C.F.O.	-6.3	14.3	35.7	-1.1	1.0	26.6	26.5	0.76
	Prontour	-6.4	12.3	35.7	-1.0	3.0	24.8	24.9	0.79
	Statistical	-7.4	16.7	25.7	0	-1.4	28.9	28.8	0.69
	Persistence	-7.5	15.0	39.9	0.1	0.3	35.2	35.1	0.61
300	Actual	-9.9	22.6	56.1					
	L.A.P.	-9.0	25.1	51.1	-0.9	-2.5	35.3	35.2	0.79
	C.F.O.	-9.9	22.0	47.9	0	0.6	36.6	36.5	0.76
	Prontour	-9.4	19.7	47.3	-0.5	2.9	36.5	36.5	0.76
	Statistical	-11.2	26.8	36.5	1.3	-4.2	40.3	40.3	0.70
	Persistence	-10.0	22.3	56.0	0.1	0.3	48.3	48.1	0.63
200	Actual	-11.2	21.5	39.1					
	L.A.P.	-11.8	23.7	42.1	0.6	-2.2	25.4	25.4	0.81
	C.F.O.	-10.1	21.8	37.7	-1.1	-0.3	24.1	24.1	0.80
	Prontour	-10.0	19.4	38.2	-1.2	2.1	25.7	25.7	0.78
	Statistical	-11.0	26.1	23.5	-0.2	-4.6	26.2	26.5	0.76
	Persistence	-11.4	21.3	39.0	0.2	0.2	30.1	30.0	0.70
100	Actual	-10.4	17.5	23.5					
	L.A.P.	-	-	-	-	-	-	-	-
	C.F.O.	-10.0	18.4	22.0	-0.4	-0.9	13.8	13.7	0.82
	Prontour	-10.9	19.9	25.1	0.5	-2.4	15.2	15.3	0.81
	Statistical	- 8.9	19.9	15.1	-1.5	-2.4	13.6	13.9	0.84
	Persistence	-10.5	17.4	23.5	0.1	0.1	15.1	15.0	0.79

Table 7. Verification of forecasts of wind at a point. January - April, 1955

(b) O.W.S. Station 'J'

Level (mb.)	Statistic	Mean component (kt.)		Standard vector deviation (kt.)	Bias in component (kt.)		Standard vector deviation of errors (kt.)	r.m.s. error (kt.)	Correlation co-efficient
		S.	W.		S.	W.			
700	Actual	5.3	7.2	27.7					
	L.A.P.	8.5	7.4	29.3	-3.2	-0.1	21.0	21.1	0.73
	C.F.O.	6.5	8.7	27.4	-1.2	-1.5	19.9	19.9	0.74
	Prontour	6.4	7.9	27.9	-1.1	-0.6	20.0	19.9	0.74
	Statistical	5.9	11.1	18.5	-0.6	-3.9	22.2	22.4	0.60
	Persistence	5.5	7.2	27.9	-0.2	0.1	25.4	25.3	0.58
500	Actual	5.5	12.3	35.5					
	L.A.P.	9.4	13.3	38.5	-3.9	-1.1	27.1	27.3	0.73
	C.F.O.	7.2	15.1	35.5	-1.7	-2.8	28.3	28.4	0.68
	Prontour	7.1	14.1	34.9	-1.6	-1.8	27.4	27.4	0.70
	Statistical	3.8	19.3	21.8	1.7	-7.0	29.3	30.0	0.57
	Persistence	5.9	12.3	35.7	-0.4	0	32.9	32.8	0.57
300	Actual	6.0	21.6	49.2					
	L.A.P.	10.3	24.0	49.9	-4.4	-2.3	37.9	38.1	0.71
	C.F.O.	8.0	21.6	46.6	-2.1	0	37.3	37.1	0.70
	Prontour	8.0	21.3	44.9	-2.1	0.3	37.4	37.3	0.69
	Statistical	3.9	26.5	30.0	2.0	-4.9	39.4	39.6	0.60
	Persistence	6.2	21.4	49.2	-0.2	0.2	46.2	46.0	0.56
200	Actual	6.7	21.4	39.9					
	L.A.P.	9.1	24.0	42.1	-2.4	-2.6	33.4	33.4	0.67
	C.F.O.	7.0	21.8	36.9	-0.2	-0.4	29.5	29.4	0.71
	Prontour	7.2	22.5	37.1	-0.5	-1.1	30.2	30.1	0.69
	Statistical	6.3	27.1	24.4	0.5	-5.7	31.3	31.7	0.62
	Persistence	7.0	21.3	40.0	-0.3	0	34.4	34.2	0.63
100	Actual	3.6	16.9	22.2					
	L.A.P.	-	-	-	-	-	-	-	-
	C.F.O.	3.8	16.9	20.9	-0.2	-0.0	14.8	14.7	0.77
	Prontour	4.5	17.9	20.6	-0.9	-0.9	14.7	14.7	0.77
	Statistical	1.9	21.1	12.9	1.7	-4.2	16.1	16.6	0.70
	Persistence	3.8	16.8	22.3	-0.2	0.2	17.6	17.5	0.69

Table 8. Levels of significance of difference between standard deviations of the forecast errors

Location	C.F.O. - Statistical		C.F.O. - L.A.P.	
	Difference (kt.)	Significance level	Difference (kt.)	Significance level
Route Shannon - Gander 500 mb.	-1.43	11%	negligible	-
Route O.W.S. 'A' - London 200 mb.	-1.85	2%	-1.24	7%
Liverpool 700 mb.	-1.58	5%	negligible	-
" 500 mb.	-2.34	4%	1.08	27%
" 300 mb.	-3.64	4%	1.36	35%
" 200 mb.	-2.07	7%	-1.27	13%
" 100 mb.	negligible	-	-	-

Table 9. Frequencies of errors greater than 50 kt., and greater than 70 kt. in point forecasts

Forecast	700 mb.		500 mb.		300 mb.		200 mb.	
	> 50 kt.	> 70 kt.	> 50 kt.	> 70 kt.	> 50 kt.	> 70 kt.	> 50 kt.	> 70 kt.
L.A.P.	2	1	12	2	35	9	16	4
C.F.O.	2	1	14	0	31	13	12	3
Prontour	2	1	10	0	35	11	12	5
Statistical	1	0	14	0	47	12	17	4
Persistence	7	0	38	6	62	33	25	7

Table 10. Large errors in point forecasts

(a) Winds forecast and observed

Date	Level (mb.)	Station	Wind observed (kt.)	Winds forecast (kt.)				
				L.A.P.	C.F.O.	Prontour	Statistical	Persistence
Jan. 15	300	Liverpool	250° 141	275° 85	300° 70	280° 72	298° 112	290° 96
" 28	200	J	190° 30*	240° 90	250° 90	255° 115	240° 51	240° 48
" 30	300	Liverpool	260° 53	170° 95	210° 45	205° 43	220° 69	200° 83
Feb. 6	300	Liverpool	320° 148	315° 50	300° 70	295° 57	290° 38	260° 18
" 7	300	J	230° 150*	245° 105	250° 75	240° 81	245° 61	235° 78*
" 14	200	J	070° 91	040° 30	010° 20	010° 16	315° 7	060° 30
" 26	500	J	310° 114	280° 60	220° 30	240° 26	280° 27	340° 39
" 26	300	J	310° 142	280° 85	220° 30	220° 43	290° 46	330° 79
" 27	300	J	250° 60	250° 130	240° 80	240° 80	285° 80	310° 142
" 27	300	Liverpool	350° 120	360° 15	300° 20	320° 23	315° 19	010° 14
March 1	300	J	310° 50	225° 80	270° 80	260° 78	220° 73	220° 114
" 2	300	J	230° 39	240° 85	220° 130	225° 108	275° 41	220° 33
" 31	200	J	250° 84	Calm	160° 35	150° 28	270° 15	260° 9

* Winds estimated

Table 10. Large errors in point forecasts

(b) Forecast errors

Date	Level (mb.)	Station	Magnitude of vector forecast error (kt.)					Comments
			L.A.P.	C.F.O.	Prontour	Statistical	Persistence	
Jan. 15	300	Liverpool	73	109	86	107	90	Position of wind maximum in a complex westerly jet stream and the location and amplitude of a short wave in the southern half of the jet, were critical for forecasting Liverpool winds.
" 28	200	J	74	80	10[illegible]	39	37	C.F.O. forecast the westward extension of a zonal jet, but in the event a major trough developed in the jet at about 27°W.
" 30	300	Liverpool	108	42	45	44	72	Forecasters underestimated the speed of translation of a short wave in the upper westerlies.
Feb. 6	300	Liverpool	96	86	99	116	139	Forecasters erred in timing the passage of a northwesterly jet on the forward side of a developing upper ridge.
" 7	300	J	55	82	71	91	72	Statistical regression towards the normal changed the wind in the wrong sense.
" 14	200	J	66	83	84	94	61	C.F.O. underestimated the development of a trough in northerly upper current. Statistical method gave a change in the wrong sense.
" 26	500	J	68	117	106	91	82	The development and subsequent eastward translation over the eastern Atlantic of a major upper ridge was not forecast by C.F.O. The prontour error for 26th at 300 mb. was the highest of the trial.
" 26	300	J	79	144	148	99	72	
" 27	300	J	70	23	23	46	124	
" 27	300	Liverpool	106	108	100	106	106	
Mar. 1	300	J	90	52	60	88	124	Eastward translation of southwesterly jet gave large persistence error.
" 2	300	J	47	91	69	31	9	Orthodox forecasters brought the main upper westerlies too far south over the Atlantic.
" 31	200	J	84	90	93	71	74	Decline of a blocking anticyclone and the development of a fresh upper ridge over the eastern Atlantic was not forecast.

The success achieved in forecasting upper winds by orthodox and statistical techniques

Appendix I: Statistical forecasts

The equation used in making the statistical forecasts was,

$$F_t - U = P(V_o - U) \qquad \ldots (A)$$

where F_t was the wind forecast for time t, V_o was the wind observed at time $t = 0$ and P and U were pre-determined constants.

Winds over the route

A statistical estimate, E, of the actual mean tailwind component over the route was obtained using the formula,

$$E - U = \sum_i a_i (v_i - u_i) \qquad \ldots (B)$$

where the V_i were winds observed at chosen stations along the route, the a_i were regression co-efficients and U and the u_i were constants. It can be shown that to minimise the forecast errors, U is best chosen to be the normal tailwind component along the route and the u_i are best chosen to be normal tailwind components of wind at the selected stations.

Substituting E for V_o in equation (A), the prediction equation is obtained in the form used in practice,

$$F_t = U + P \sum_i a_i (v_i - u_i) \qquad \ldots (C)$$

Values of U and u_i are given in Table A. Table B gives the bearings of the routes at selected stations. Regression co-efficients: a_i for the estimated actuals; Pa_i for the forecasts; are contained in Table C.

To obtain E for 1500 G.M.T. in the absence of a 1500 G.M.T. observation, a geostrophic estimate, made from the 1500 G.M.T. chart, was used in equation B.

For the forecasts, the winds u_i were normally 2100 G.M.T. observations. In the absence of a 2100 G.M.T. observation the previous 1500 G.M.T. observation was used. If neither a 2100 G.M.T. nor a 1500 G.M.T. observation was available then a geostrophic estimate was made from the 1500 G.M.T. contour chart. For the Irish and Canadian stations (route Shannon - Gander) the rules were:

(i) if Valentia is not available at 2100 G.M.T.: use the Aldergrove wind; but if neither is available at 2100 G.M.T.: use the mean of the two 1500 G.M.T. winds;

(ii) for Newfoundland, either Argentia or Stephenville can be used, but if both are received for the same hour, the mean of the two is taken.

Winds at a point

The point forecasts were made using equation (A). Forecasts were normally based on the 2100 G.M.T. observed wind, when P was taken to be 0.65. If a 2100 G.M.T. observation was missing, the previous 1500 G.M.T. observation was used, when P was taken as 0.55. In the absence of both observations a geostrophic estimate was made from the 1500 G.M.T. chart. Table D contains values of U.

Table A

Normal tailwind components along track

(a) Route O.W.S. 'A' - London 200 mb.

Month	u_i in kt.					U in kt.
	O.W.S. 'A'	O.W.S. 'I'	Aldergrove	Liverpool	Crawley	
January	32	32	33	33	33	33
February	29	30	30	29	28	30
March	26	29	27	25	23	27
April	23	28	24	21	18	24

(b) Route Shannon - Gander 500 mb.

Month	u_i in kt.				U in kt.
	Valentia/Aldergrove	O.W.S. 'J'	O.W.S. 'C'	Stephenville/Argentia	
January	-30	-31	-37	-41	-35
February	-27	-28	-33	-39	-32
March	-24	-25	-29	-36	-28
April	-21	-22	-25	-34	-25

Table B

Bearings of the routes at selection stations

	O.W.S. 'A'	O.W.S. 'I'	Aldergrove	Liverpool	Crawley
O.W.S. 'A' - London	110°	120°	130°	135°	135°
Shannon-Gander	Gander 240°	O.W.S. 'C' 260°	O.W.S. 'J' 270°	Shannon 285°	

Table C

(a) Regression co-efficients for route Ship 'A' - London 200 mb.

	O.W.S. 'A'	O.W.S. 'I'	Aldergrove	Liverpool	Crawley
Actual (all observations 1500Z)	0.15	0.39	0.22	0.16	0.05
Forecast (all observations 2100Z)	0.12	0.32	0.17	0.13	0.04
F/c (O.W.S. 'A' 1500Z, others 2100Z)	0.10	0.34	0.18	0.13	0.04
F/c (O.W.S. 'I' 1500Z, others 2100Z)	0.18	0.23	0.23	0.13	0.03
F/c (Aldergrove 1500Z, others 2100Z)	0.12	0.34	0.08	0.24	0.03
F/c (Liverpool 1500Z, others 2100Z)	0.12	0.32	0.23	0.04	0.09
F/c (Crawley 1500Z, others 2100Z)	0.12	0.32	0.17	0.15	0.03
F/c (O.W.S. 'A' and 'I' 1500Z, others 2100Z)	0.09	0.30	0.24	0.13	0.03

Table C

(b) Regression co-efficients for route Shannon - Gander 500 mb.

	Valentia/ Aldergrove	O.W.S. 'J'	O.W.S. 'C'	Stephenville/ Argentia
Actual (all observations 1500Z)	0.12	0.22	0.39	0.20
F/c (all observations 2100Z)	0.09	0.17	0.29	0.15
F/c (Val./Ald. 1500Z, others 2100Z)	0.07	0.18	0.28	0.15
F/c (O.W.S. 'J' 1500Z, others 2100Z)	0.11	0.13	0.30	0.15
F/c (O.W.S. 'C' 1500Z, others 2100Z)	0.07	0.19	0.25	0.15
F/c (Ste./Arg. 1500Z, others 2100Z)	0.09	0.16	0.29	0.13
F/c (Val./Ald. O.W.S. 'J' 1500Z, others 2100Z)	0.08	0.14	0.30	0.15
F/c (Val./Ald. O.W.S. 'C' 1500Z, others 2100Z)	0.06	0.20	0.23	0.15
F/c (Val./Ald. Ste./Arg. 1500Z, others 2100Z)	0.07	0.18	0.28	0.13
F/c (O.W.S. 'J', O.W.S. 'C' 1500Z, others 2100Z)	0.10	0.14	0.25	0.15
F/c (O.W.S. 'J', Ste./Arg. 1500Z, others 2100Z)	0.11	0.13	0.30	0.13
F/c (O.W.S. 'C', Ste./Arg. 1500Z, others 2100Z)	0.08	0.19	0.25	0.13
F/c (Val./Ald. 'J', 'C' 1500Z, others 2100Z)	0.08	0.14	0.25	0.15
F/c (Val./Ald. 'J', Ste./Arg. 1500Z, others 2100Z)	0.08	0.14	0.30	0.13
F/c (Val./Ald. 'C', Ste./Arg. 1500Z, others 2100Z)	0.06	0.20	0.24	0.13
F/c ('J', 'C', Ste./Arg. 1500Z, others 2100Z)	0.10	0.14	0.25	0.13

Table D

Normal values*, U, of wind at O.W.S. 'J' and Liverpool

Height (mb.)	O.W.S. 'J'								Liverpool							
	Jan.		Feb.		Mar.		Apr.		Jan.		Feb.		Mar.		Apr.	
	deg.	kt.	deg.	kt.	deg.	kt.	deg.	kt.	deg.	kt.	deg.	kt.	deg.	kt.	deg.	kt.
100	265	35	270	30	270	25	275	20	280	30	280	27	275	23	275	20
200	260	48	260	43	265	38	265	32	295	40	290	37	285	33	270	30
300	260	48	270	43	270	38	280	32	290	41	285	38	285	34	280	31
500	270	37	270	33	270	28	270	23	280	30	280	27	285	24	285	21
700	250	22	250	19	250	16	250	13	270	15	270	14	270	12	170	11

* These values were obtained from Reference 10

Appendix II: Significance tests

The series of forecast errors discussed in this paper contain small but significant serial correlations, and furthermore the errors of one forecasting technique are correlated with the errors of each of the other forecasting techniques. The usual criteria of statistical significance, which depend upon assumptions of independence of the samples, do not therefore apply.

A device commonly adopted in dealing with a serially correlated set of observations is to estimate the number of independent observations to which they are equivalent. Lewis and McIntosh (2) compare the standard error of the mean of n independent observations, namely $\sigma/\sqrt{n}$, where σ is the population standard deviation, with the standard error of the mean of n serially correlated observations which is,

$$\sigma\left[1+\frac{2}{n}\left\{(n-1)p_1+(n-2)p_2+\cdots+p_{n-1}\right\}\right]^{\frac{1}{2}}\Big/\sqrt{n} \quad \cdots \text{(i)}$$

where p_1 is the correlation between consecutive terms, p_2 is the correlation between alternate terms, p_3 is the correlation between every third term, and so on. They suggest that the number, n', of effectively independent terms be reckoned for statistical purposes as $n\Big/\left[1+\frac{2}{n}\left\{(n-1)p_1+(n-2)p_2+\cdots+p_{n-1}\right\}\right]$. The standard error of the persistent series can then be written as $\sigma/\sqrt{n'}$. Strictly, the use of an n' so obtained, in computing standard errors, is only justified rigorously when mean values are to be dealt with, but it also provides a working basis for assessing the significance of other statistical parameters such as the standard deviation.

By analogy with the conception of an "equivalent number of independent observations" there can be defined "an equivalent number of independent comparisons" for use when two series which are correlated with each other have to be compared. Lewis and McIntosh (3) give $\left[2\sigma^2(1-r)/n\right]^{\frac{1}{2}}$ as the standard error of the difference between the means of two series, of n terms, which are related to each other with co-efficient of correlation r, but which individually have no serial correlation. The two series are assumed to have a common standard deviation σ. In this case an equivalent number of independent comparisons, n'', can be defined equal to $n/(1-r)$. The standard error of the difference between the means is then $(2\sigma^2/n'')^{\frac{1}{2}}$. If series to be dealt with contain serial correlations, then as a first step the equivalent number of independent observations can be obtained for each one; but difficulties then arise in obtaining n'' if n' is not the same for each series. It is therefore necessary to derive an expression for the standard error of the difference between the means ab initio.

Given two series, $x_1 \ldots\ldots\ldots\ldots x_n$ and $y_1 \ldots\ldots\ldots\ldots y_n$ with correlation r between x and y, and serial correlations p_1, p_2 etc. in the x and q_1, q_2 etc. in the y. Let $\bar{x}$, $\bar{y}$ be the sample means and X, Y be population means for x and y. Let $t_{i,j}$ be the correlation between x_i and y_j.

Define

$$P = \frac{1}{n}\left\{(n-1)p_1+(n-2)p_2+\cdots+p_{n-1}\right\};$$

$$Q = \frac{1}{n}\left\{(n-1)q_1+(n-2)q_2+\cdots+q_{n-1}\right\};$$

$$T = \frac{1}{n}\left\{(n-1)t_{1,2}+(n-2)t_{1,3}+\cdots+t_{1,n}\right\}.$$

/Then,

Then, $\left[\text{standard error of } (\bar{x}-\bar{y})\right]^2 = \frac{1}{m}\sum\left[(\bar{x}-\bar{y})-(X-Y)\right]^2$,

the sum being taken over m samples,

$$= \frac{1}{m}\sum_m (\bar{x}-X)^2 + \frac{1}{m}\sum_m (\bar{y}-Y)^2 - \frac{2}{m}\sum_m (\bar{x}-X)(\bar{y}-Y) \quad \ldots \text{(ii)}$$

The first two terms of (ii) are the squares of the standard errors of the sample means and by (i) are equal to $\sigma_x^2(1+2P)/n$ and $\sigma_y^2(1+2Q)/n$ respectively. The third term of (ii) may be expanded as,

$$-\frac{2}{m n^2}\sum_m \left[(x_1-X)+(x_2-X)+\cdots+(x_n-X)\right]\left[(y_1-Y)+\cdots+(y_n-Y)\right],$$

which when m is large, reduces to,

$$-\frac{2}{n^2}\left[n r \sigma_x\sigma_y + 2(n-1)\sigma_x\sigma_y t_{1,2} + 2(n-2)\sigma_x\sigma_y t_{1,3} + \cdots + 2 t_{1,n}\right]$$

(assuming that $t_{i,j} = t_{j,i}$), which expression may be written as

$$-2\frac{\sigma_x\sigma_y}{n}(r+2T).$$

Equation (ii) may then be written,

$$\left[\text{s.e.}(\bar{x}-\bar{y})\right]^2 = \frac{\sigma_x^2}{n}\left[1+2P\right] + \frac{\sigma_y^2}{n}\left[1+2Q\right] - \frac{2\sigma_x\sigma_y}{n}(r+2T)$$

$$= \frac{\sigma_x^2+\sigma_y^2}{n}\left[1+\frac{2\left\{P\sigma_x^2+Q\sigma_y^2-\sigma_x\sigma_y(r+2T)\right\}}{\sigma_x^2+\sigma_y^2}\right]$$

$$= (\sigma_x^2+\sigma_y^2)/n''$$

where
$$n'' = n\Big/\left[1+\frac{2\left\{P\sigma_x^2+Q\sigma_y^2-\sigma_x\sigma_y(r+2T)\right\}}{\sigma_x^2+\sigma_y^2}\right] \quad \ldots \text{(iii)}$$

In the case where there are no serial correlations and r = 0

$$\left[\text{s.e.}(\bar{x}-\bar{y})\right]^2 = (\sigma_x^2+\sigma_y^2)/n$$

and so n'' can be defined as the equivalent number of independent comparisons. Table (i) gives values of serial correlations and correlations between the variates, for forecast errors of wind over the routes and of wind at the points. For the points the correlations are stretch vector correlations. A serial correlation greater than .18 is inconsistent at the 5% level with an assumption that the forecast errors are uncorrelated. The serial correlations quoted are the correlations between consecutive terms (i.e., p_1 or q_1):

/in

in the work which follows it has been assumed that $p_2 = p_1^2$, $p_3 = p_1^3$ etc. There are theoretical grounds for making this assumption (see for example Brooks and Carruthers (4)), and some practical justification for expecting the assumption to be valid for wind statistics may be found in a study of the temporal correlations given by Durst (1). In this work, the p_1's were usually so small that it was not necessary to take into account correlations of a higher order than p_2.

Values of $t_{i,j}$ were not computed from the variates themselves: they were derived from the known correlations p, q, r. $t_{1,2}$ was obtained as follows. With the usual notation for partial correlations,

$$r_{x_1y_2 \cdot x_2y_1} = \frac{\left(r_{x_1y_2 \cdot y_1} - r_{x_1x_2 \cdot y_1} r_{y_2x_2 \cdot y_1}\right)}{\sqrt{\left(1 - r^2_{x_1x_2 \cdot y_1}\right)\left(1 - r^2_{y_2x_2 \cdot y_1}\right)}}$$

since x_1 and y_2 are related only through x_2 and y_1, $r_{x_1y_2 \cdot x_2y_1}$ can be taken to be zero and so,

$$r_{x_1y_2 \cdot y_1} = r_{x_1x_2 \cdot y_1} r_{y_2x_2 \cdot y_1}$$

Expanding the last relation in terms of the total correlations

$$(t_{1,2} - rq)(1 - t_{1,2}^2) = (p - r t_{1,2})(r - q t_{1,2})$$

when $t_{1,2}$ is taken equal to $t_{2,1}$.

The required $t_{1,2}$ is a root of the cubic

$$t_{1,2}^3 - (pq + r^2 + 1)t_{1,2} + r(p + q) = 0 .$$

In the present work, the values of p, q and r were such that all the roots of this cubic were real, but two of the roots were numerically greater than one, leaving only one physically meaningful solution. $t_{1,3}$ can be obtained as a root of the cubic

$$t_{1,3}^3 - (p^2q^2 + r^2 + 1)t_{1,3} + r(p^2 + q^2) = 0, \text{ and so on.}$$

In assessing the significance of the correlations it is helpful to consider how they can arise in the case of statistical and persistence forecasts.
Equation (9), Section 3 of the main paper gives the formula for the serial correlation in the statistical forecast errors. If it is assumed that there is no bias in the forecasts ($U = \overline{V}$) and that P has been chosen approximately equal to the serial correlation, R_t, then the expression (9) reduces to $R_t\left(R_t^2 - R_{2t}\right)/\left(1 - R_t^2\right)$ which vanishes when $R_t^2 = R_{2t}$. ($P_2 = P_1^2$ in the notation of the previous paragraph). It follows that there will be little or no serial correlation in the statistical forecast errors unless they contain a bias: this explains why the statistical forecast error correlation

/given

given in Table (i) is appreciable for the 500 mb. route but is negligible for the 200 mb. route. Equation (10) is the expression for the persistence forecast errors. Substituting values of R_{24} and R_{48} quoted by Durst (1) for Larkhill in expression (10), values for the serial correlation in persistence forecasts of -.21 and -.20 are obtained for winds at 200 mb. and 100 mb. respectively. These are comparable with correlations of -.16 and -.21 given for Liverpool in Table (i).

Route winds

Significance levels of the bias in the C.F.O. and L.A.P. 200 mb. forecasts, given in Section 4 of the main paper, were assessed using expression (i) of this appendix.

The standard error of the difference between the standard deviations of two normally distributed variates is $[(\sigma_1^2/n_1 + \sigma_2^2/n_2)/2]^{\frac{1}{2}}$. To obtain the levels of significance given in Table 8 for the routes, n_1 and n_2 were taken equal to the appropriate n'', which was itself calculated from equation (iii) above.

Point winds

For the points, the forecast errors are vector errors. Brooks and Carruthers (5) state that the observed winds conform to a normal circular distribution, and it can be expected that the forecasts, both statistical and orthodox, will also conform to this distribution: but it does not follow that the forecast errors should. It can be shown that the standard deviation S_x of the x components of the persistence forecast errors is given by $S_x^2 = 2\sigma_x^2(1-R_x)$ where R_x is the serial correlation in the x components of the observed winds. So $S_w \neq S_s$, the subscripts referring to west and south components, unless $R_w = R_s$ and $\sigma_w = \sigma_s$. Calculations were made of R_w and R_s for the Liverpool winds at 500 mb. and 200 mb. At 500 mb. R_w was .67 and R_s was .55, the difference between the two being not significant. At 200 mb. R_w was .85 (lower 5% confidence limit 0.77), and R_s was .54 (upper 5% confidence limit .68): here the difference was clearly significant.

In view of the considerations of the previous paragraph, tests were made for circularity in the distributions of each type of forecast error at 500 mb. and 200 mb. The test used was that developed by Mauchly and quoted in Brooks and Carruthers (6). The values of the ellipticity statistic which were obtained were not significant at the 5% level for statistical, persistence and orthodox forecasts at 500 mb., but values obtained for the 200 mb. forecasts were all significant at the 5% level.

The elliptic nature of the distribution of 200 mb. errors is evidently due largely to an interesting property of the wind flow at high levels, namely that the west component of the wind varies more slowly with time than does the south component. This feature has been noticed previously in connexion with 100 mb. wind statistics, e.g. in calculating the standard errors of 6-hr. variations of west and south components of 100 mb. wind, it was found that the variance of the west component variations was smaller than the variance of the south component variations (7). For present purposes it has been assumed that the non-circularity occurs only in the lower-stratospheric regime, and that the tropospheric forecast errors conform to a normal circular distribution.

For a normal circular distribution of vectors it is convenient to sum the departures from the vector mean over all directions to obtain a scalar variate V. The equation to the frequency curve of V is, $f(v) = \frac{2N}{\sigma^2} v e^{-v^2/\sigma^2}$ (Brooks and Carruthers (8)).

The second moment, μ_2, or variance of this distribution is given by $\int_0^\infty \frac{2}{\sigma^2} v^3 e^{-v^2/\sigma^2}$ and the fourth moment, μ_4, is $\int_0^\infty \frac{2}{\sigma^2} v^5 e^{-v^2/\sigma^2}$.

/Integration

Integration yields $\mu_2 = \sigma^2$ and $\mu_4 = 2\sigma^4$. This distribution is obviously not normal. The standard error of the difference between the standard deviations of two such distributions may be obtained using the formula,

$$\left[\text{s.e. of } (\sigma_1 - \sigma_2)\right]^2 = \frac{\mu_4 - \mu_2^2}{4\mu_2 n_1} + \frac{\nu_4 - \nu_2^2}{4\nu_2 n_2} \quad \ldots\ldots \text{(Yule and Kendall (9))}$$

$$= \frac{1}{4}\left\{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right\} \quad \ldots \text{(iv)}$$

The significance levels quoted in Table 8 for Liverpool forecasts at 700 mb., 500 mb. and 300 mb. have been calculated using equation (iv), where $n_1 = n_2 = n''$, obtained from equation (iii).

For Liverpool 200 mb. errors, the south and west components were treated as two separate normal distributions. The correlations between south and west component errors were: .09 for persistence, .12 for C.F.O., and .16 for statistical forecasts (5% significance level was .22). The probabilities, p_i and q_i, of occurrence of the observed differences between relevant pairs of standard deviations of forecast errors, were computed for each component. The values given in Table 8 for 200 mb. are the probabilities, $p_i q_i$, of simultaneous occurrence of the observed differences of standard deviation in the two components together, the two component distributions being treated as mutually independent.

Tests for the representativeness of the 1955 data

The percentage frequency with which deviations from the normal of sample mean winds, equal to or greater than the deviation which occurred for the trial period, can be expected to occur, are given in Table 4 of the paper, in the column headed "Frequency of 1955 departure". To obtain these values, (a) the 1948-55 means for January to April were taken as the normals, and (b) the standard error of the 1955 mean was calculated using a form of equation (i) derived by Brooks, viz.,

$$\left[\text{s.e. mean}\right]^2 = \frac{\sigma^2}{n}\left\{1 + \frac{2R_t}{1-R_t}\left(1 - \frac{1}{n}\cdot\frac{1-R_t^n}{1-R_t}\right)\right\}$$

This expression is derived assuming that $R_{nt} = R_t^n$. As previously mentioned, this assumption is probably reasonable in the case of wind data where t = 24 hr. for small n, i.e., n = 2 or 3, but in the absence of practical confirmation, it becomes increasingly dubious as n increases. Steps (a) and (b) are both by no means satisfactory, but are made in the absence of further data. Consequently, as mentioned in Section 5 of the paper, the percentage frequencies given in Table 4 must be treated with some reserve, but they do give a guide to the normality of the trial period.

/Table (i)

Table (i). Correlations in forecast errors

(a) Serial correlations in forecast errors

Forecast error	Routes		Liverpool *	
	500 mb.	200 mb.	500 mb.	200 mb.
L.A.P.	.16	-.01	-.08	.11
C.F.O.	.22	.08	-.00	.15
Statistical	.23	.04	-.05	.22
Persistence	.03	-.13	-.21	-.16

(b) Correlations between forecast errors

Forecast errors	Routes		Liverpool *	
	500 mb.	200 mb.	500 mb.	200 mb.
C.F.O. v. Statistical	.47	.44	.58	.54
C.F.O. v. L.A.P.	.61	.58	.66	.53

* Liverpool correlations are stretch vector correlation co-efficients

Fig. 1

Frequencies of forecast errors — route forecasts

January - April 1955

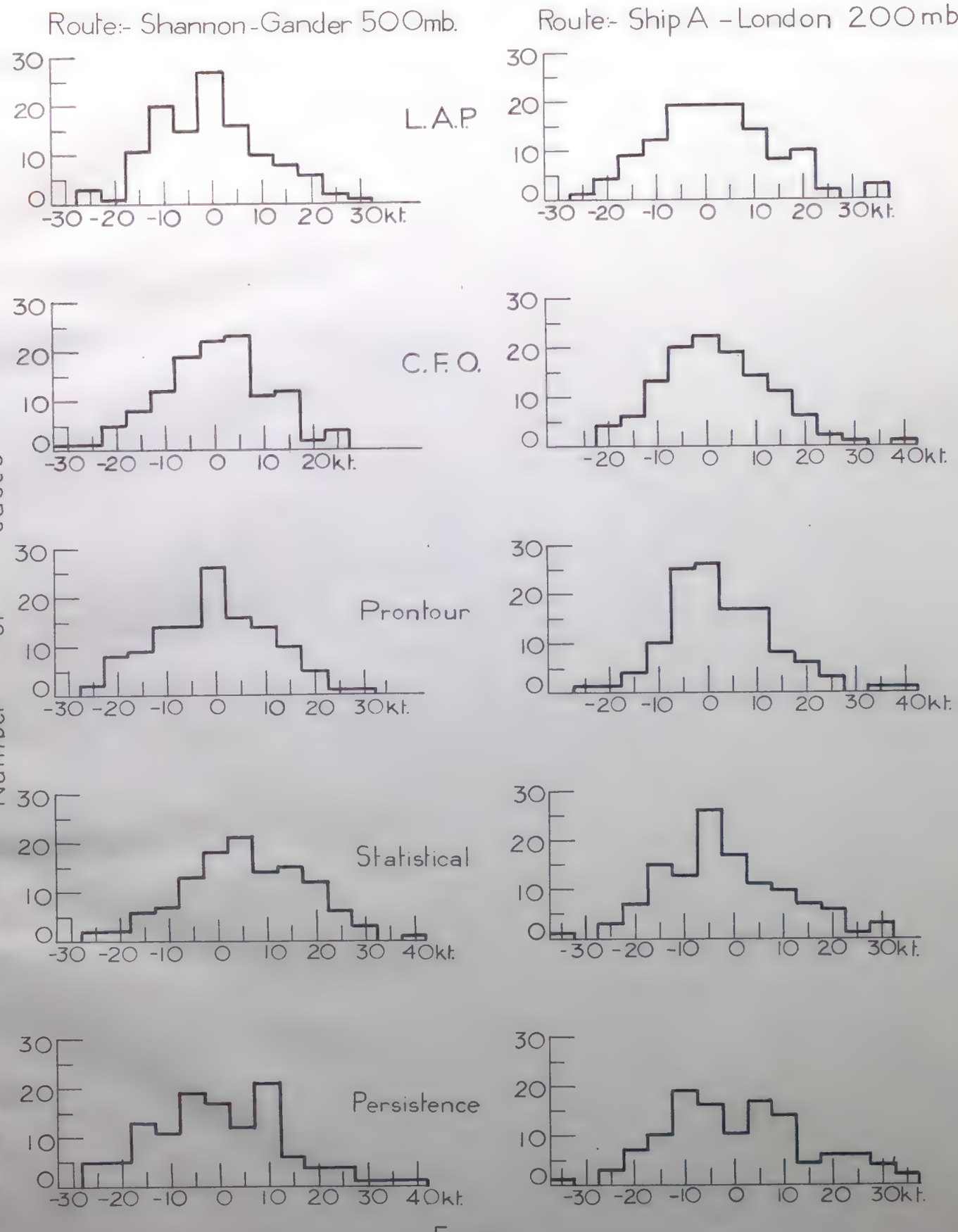


Fig. 2(a)

Individual forecasts for route Shannon-Gander
500 mb. — January, 1955

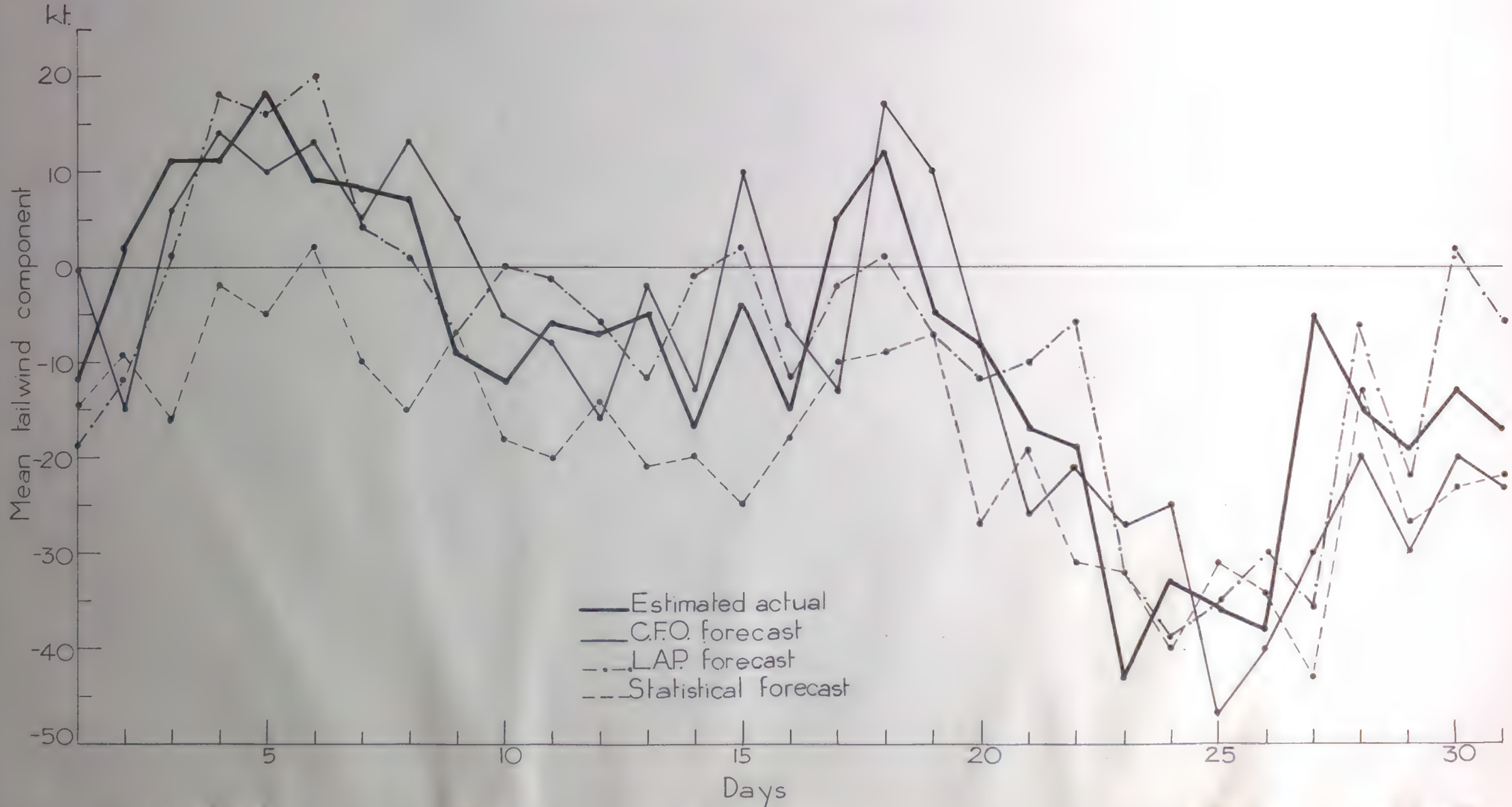


Fig. 2 (b)

Individual forecasts for route Shannon-Gander

500 mb. — February 1955

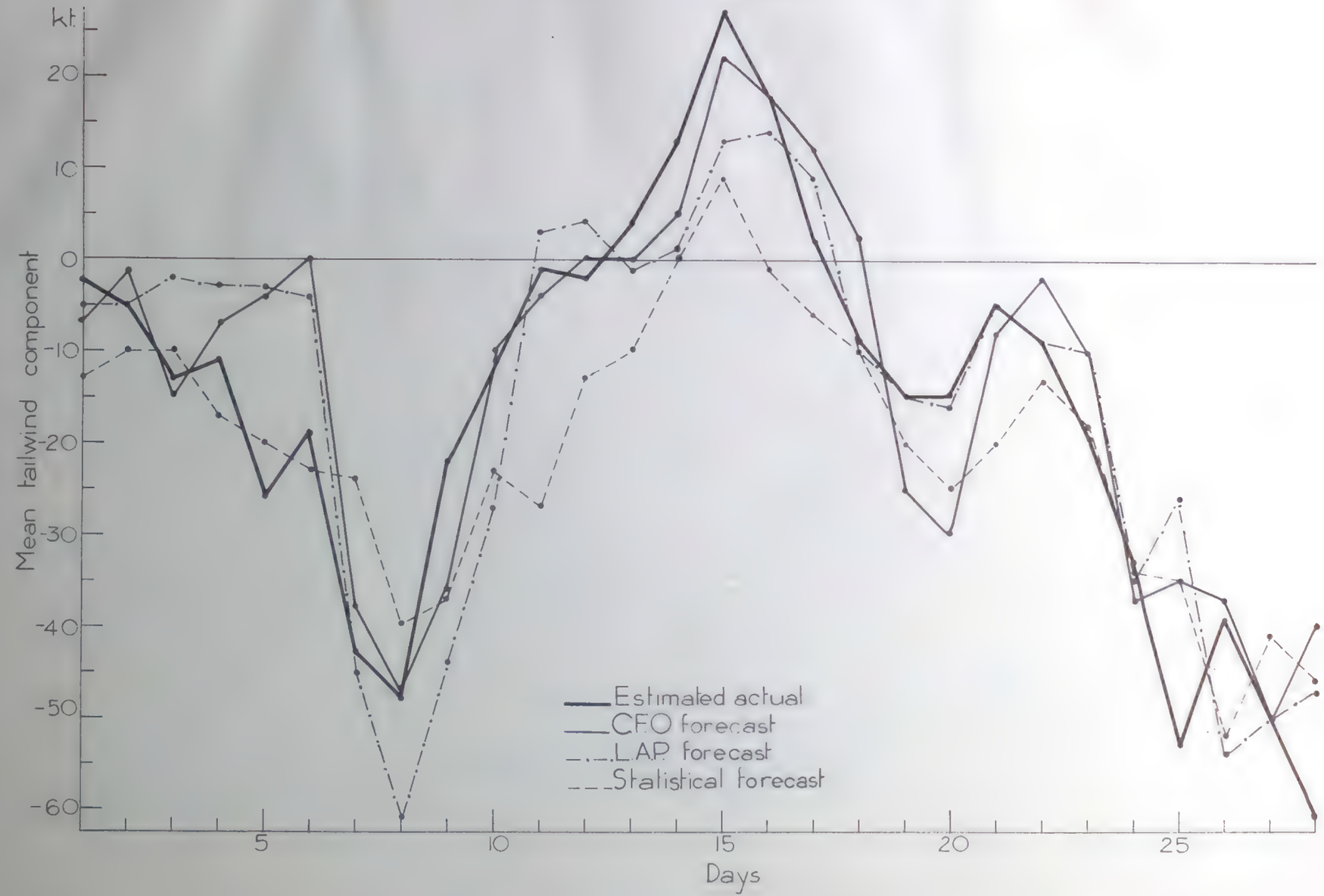


Fig. 2(c)

Individual forecasts for route Shannon-Gander
500 mb. — March, 1955

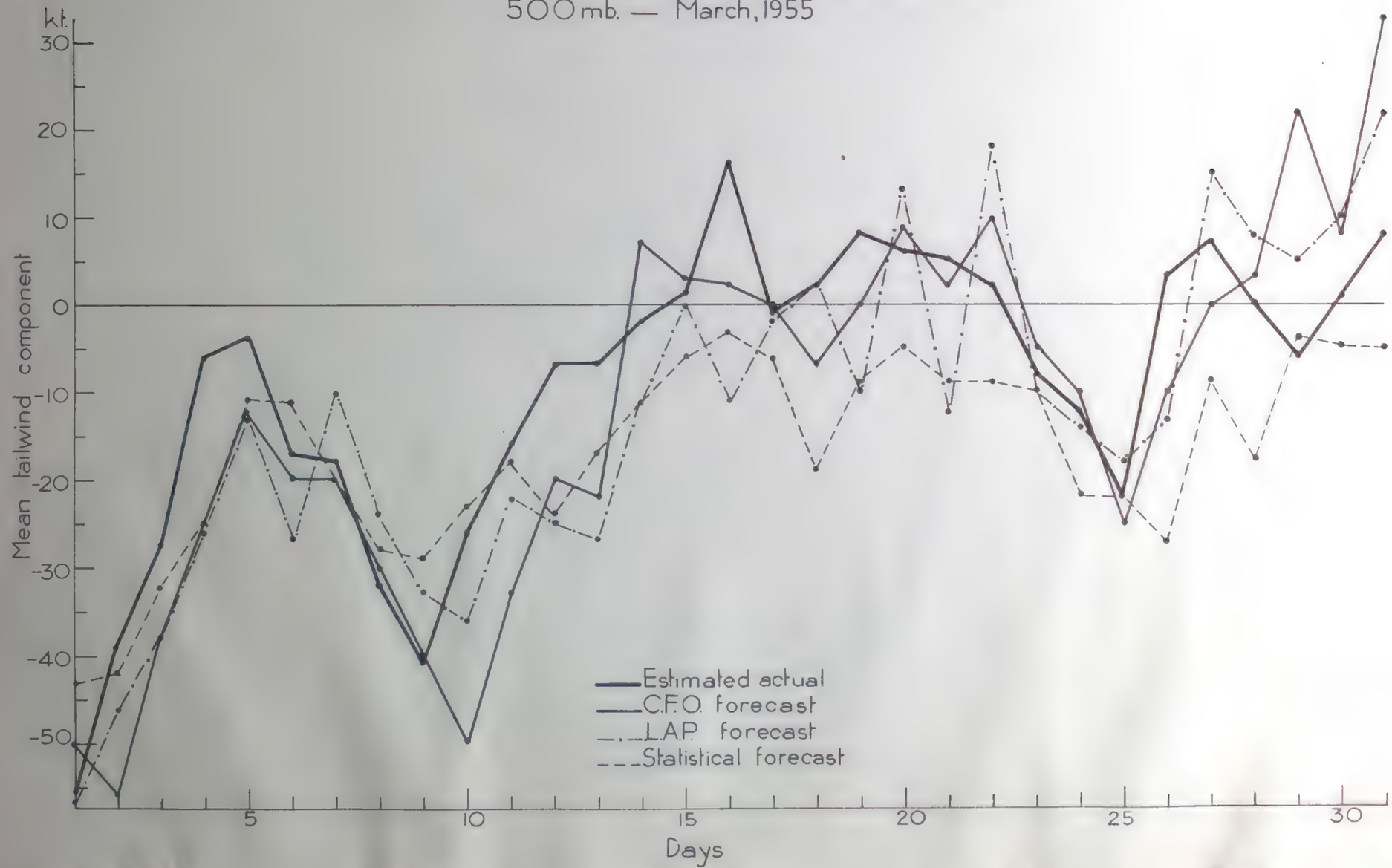
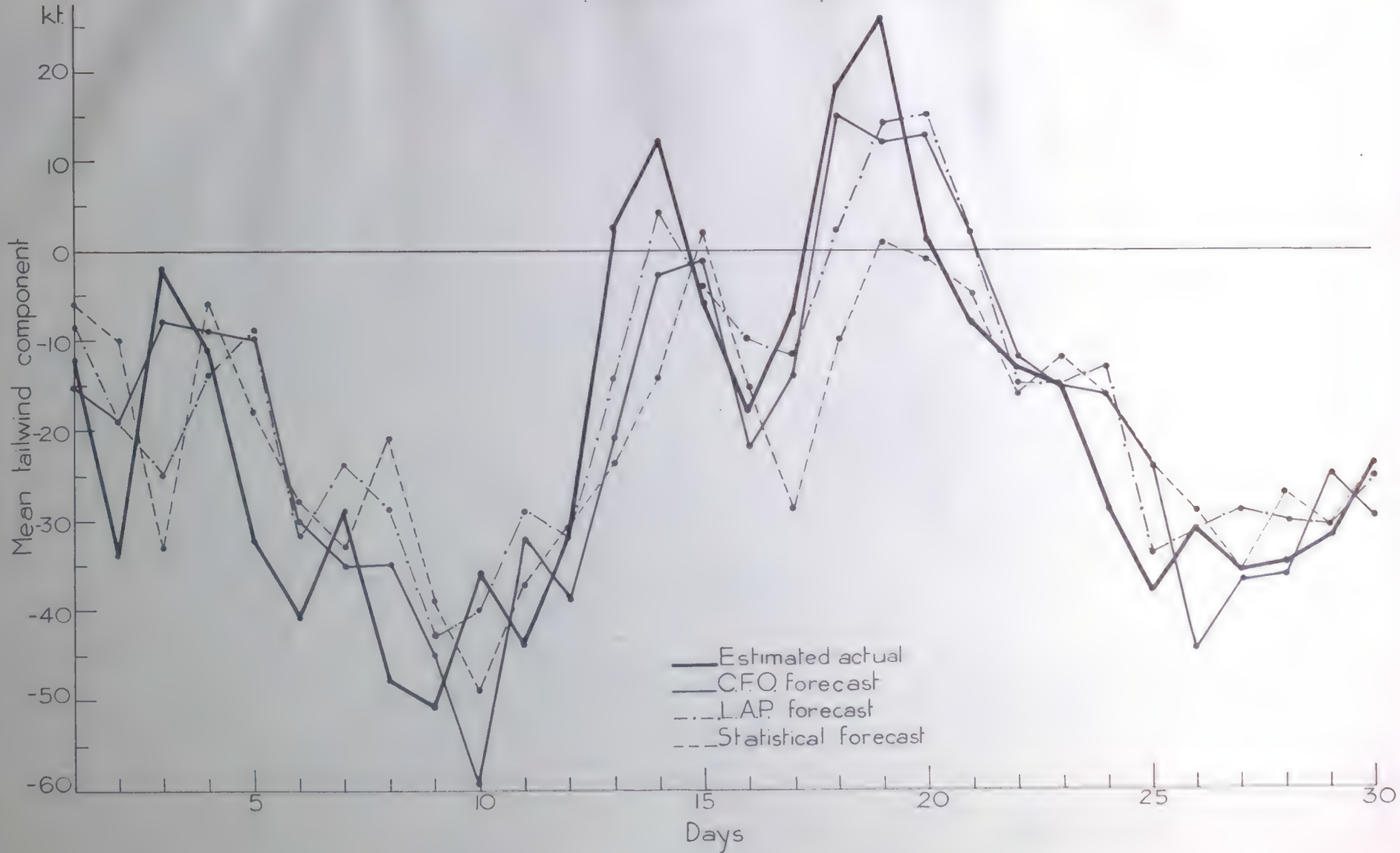


Fig.2(d)

Individual forecasts for route Shannon-Gander
500mb. — April, 1955



M.R.P. 953

Fig. 3(a)

Individual forecasts for route O.W.S. 'A'-London
200mb. —January, 1955

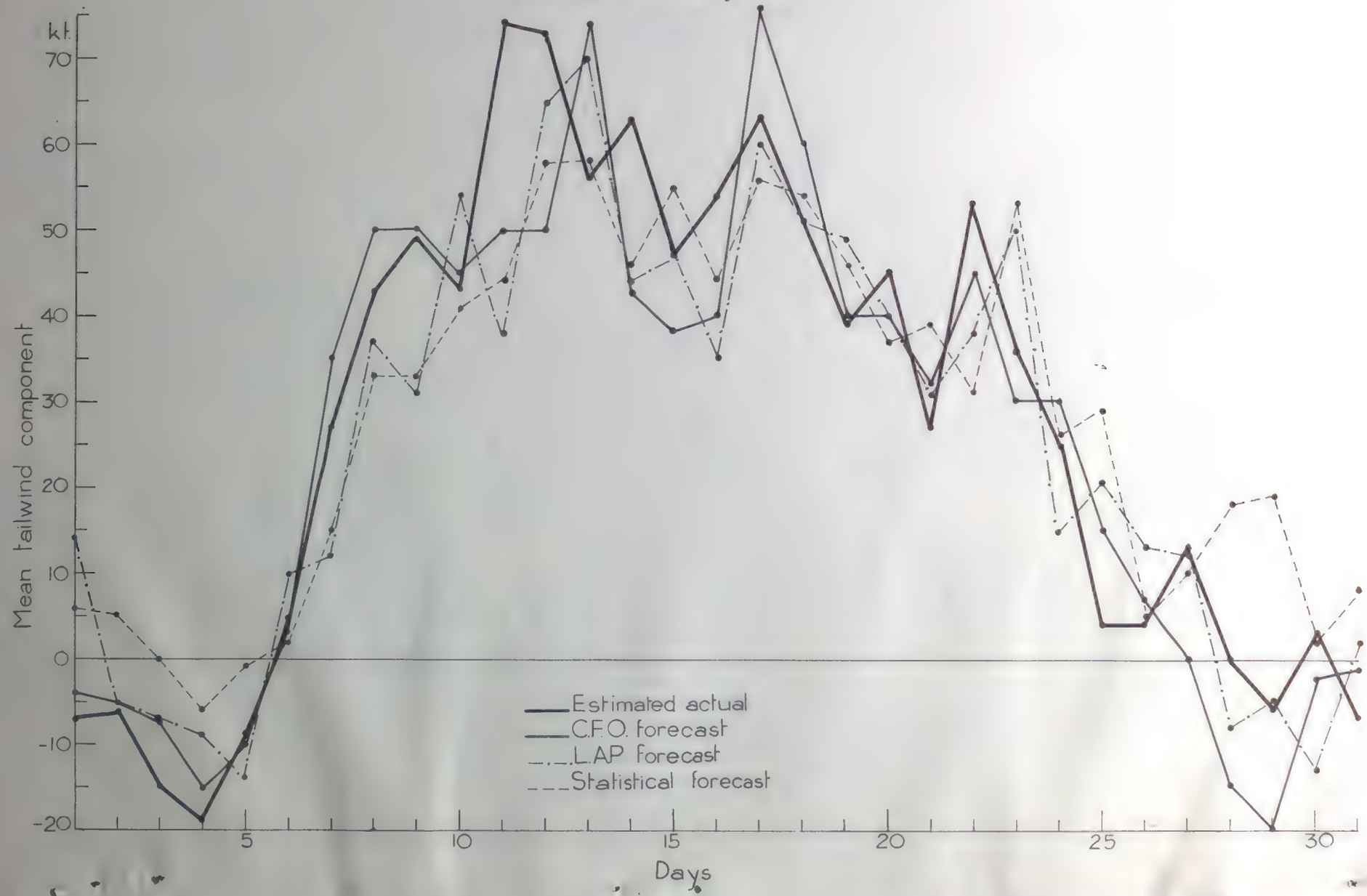


Fig. 3(b)

Individual forecasts for route O.W.S. "A" – London
200 mb. — February 1955

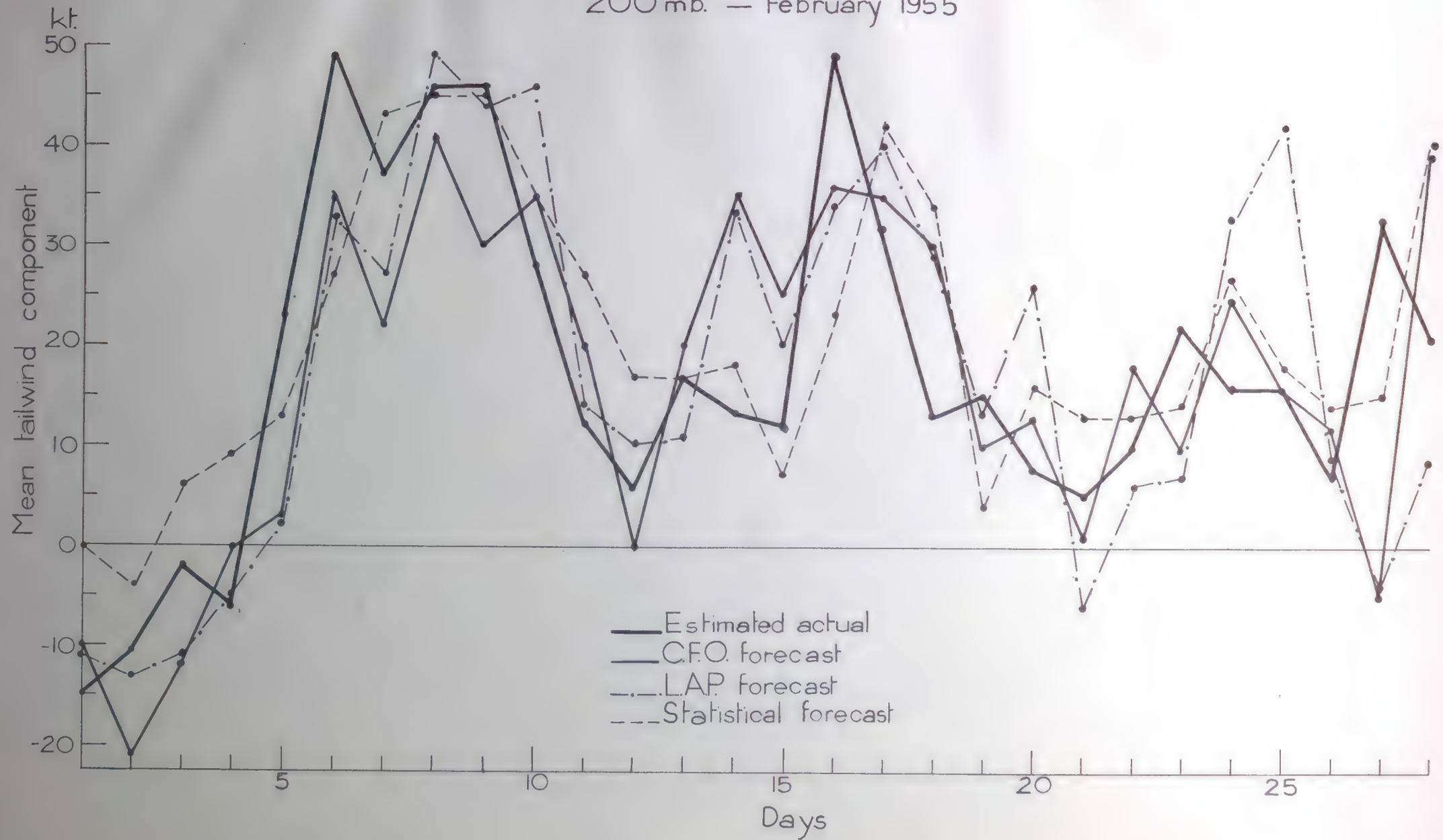


Fig. 3(c)

Individual forecasts for route O.W.S.A – London
200mb. — March, 1955

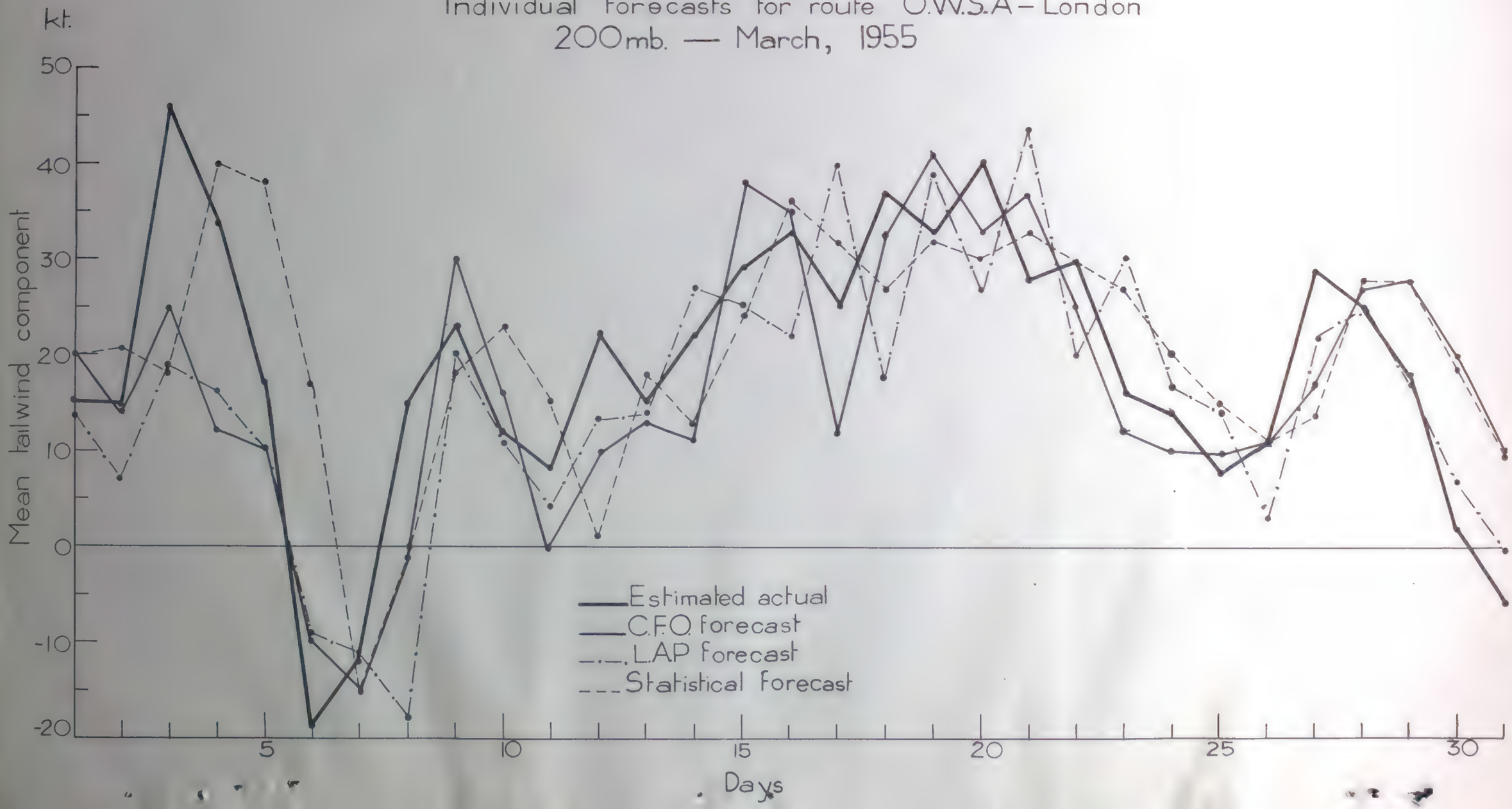


Fig. 3(d)

Individual forecasts for route O.W.S. "A" – London
200 mb. – April, 1955

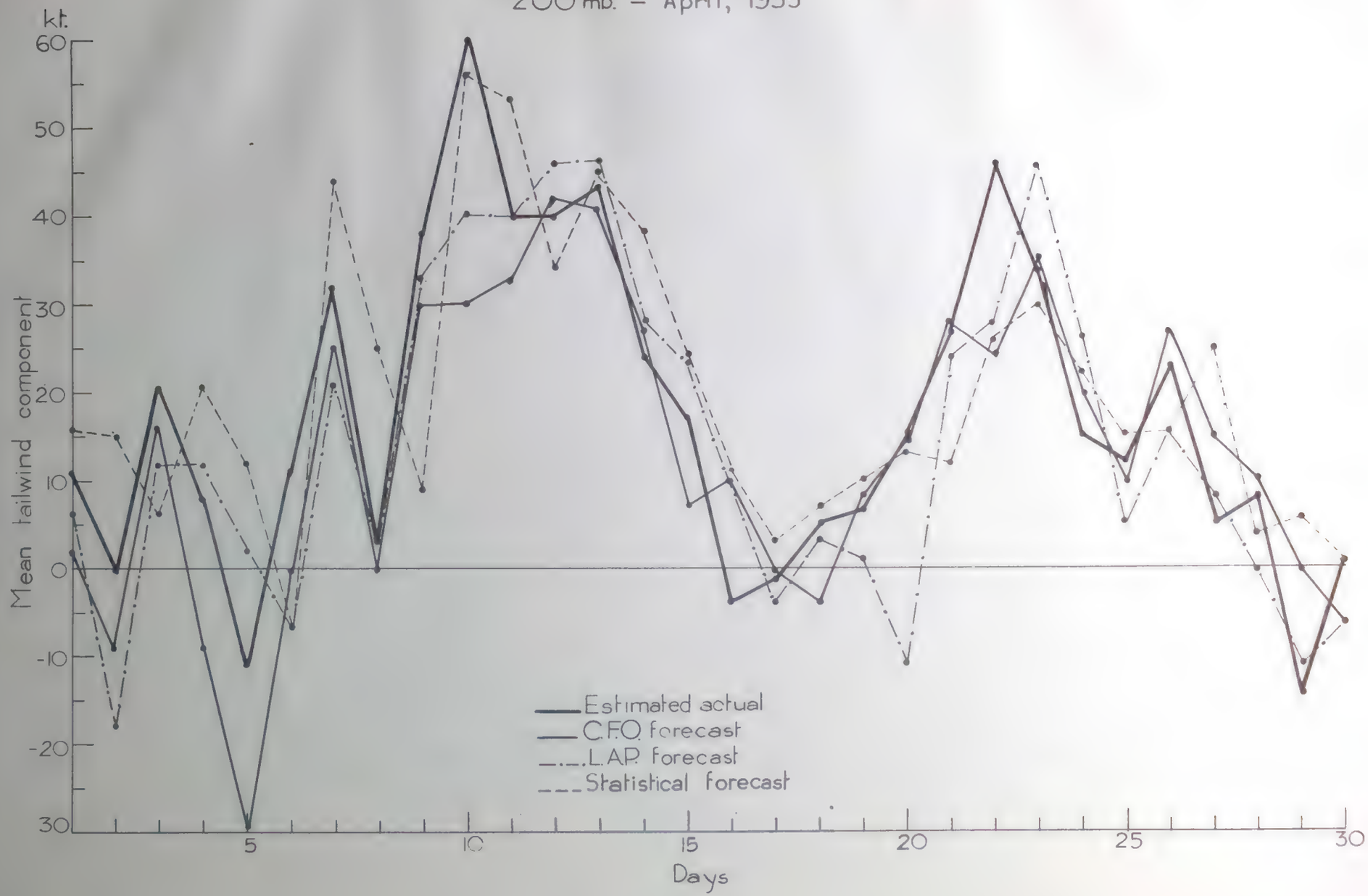


Fig. 4(a)

Frequencies of forecast errors – point forecasts, 700 mb.
January–April 1955

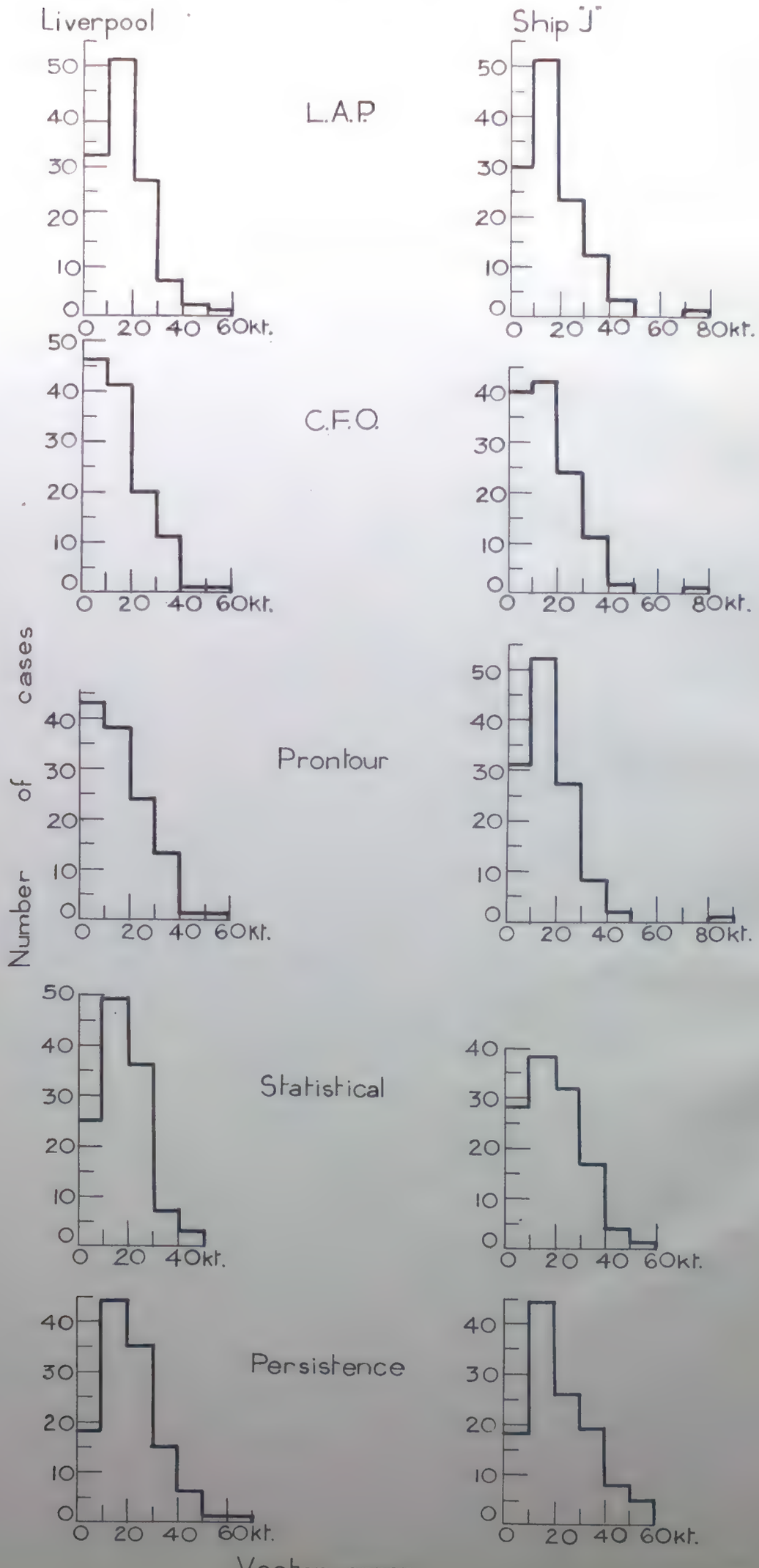


Fig. 4 (b)

Frequencies of forecast errors – point forecasts, 500 mb.
January – April, 1955

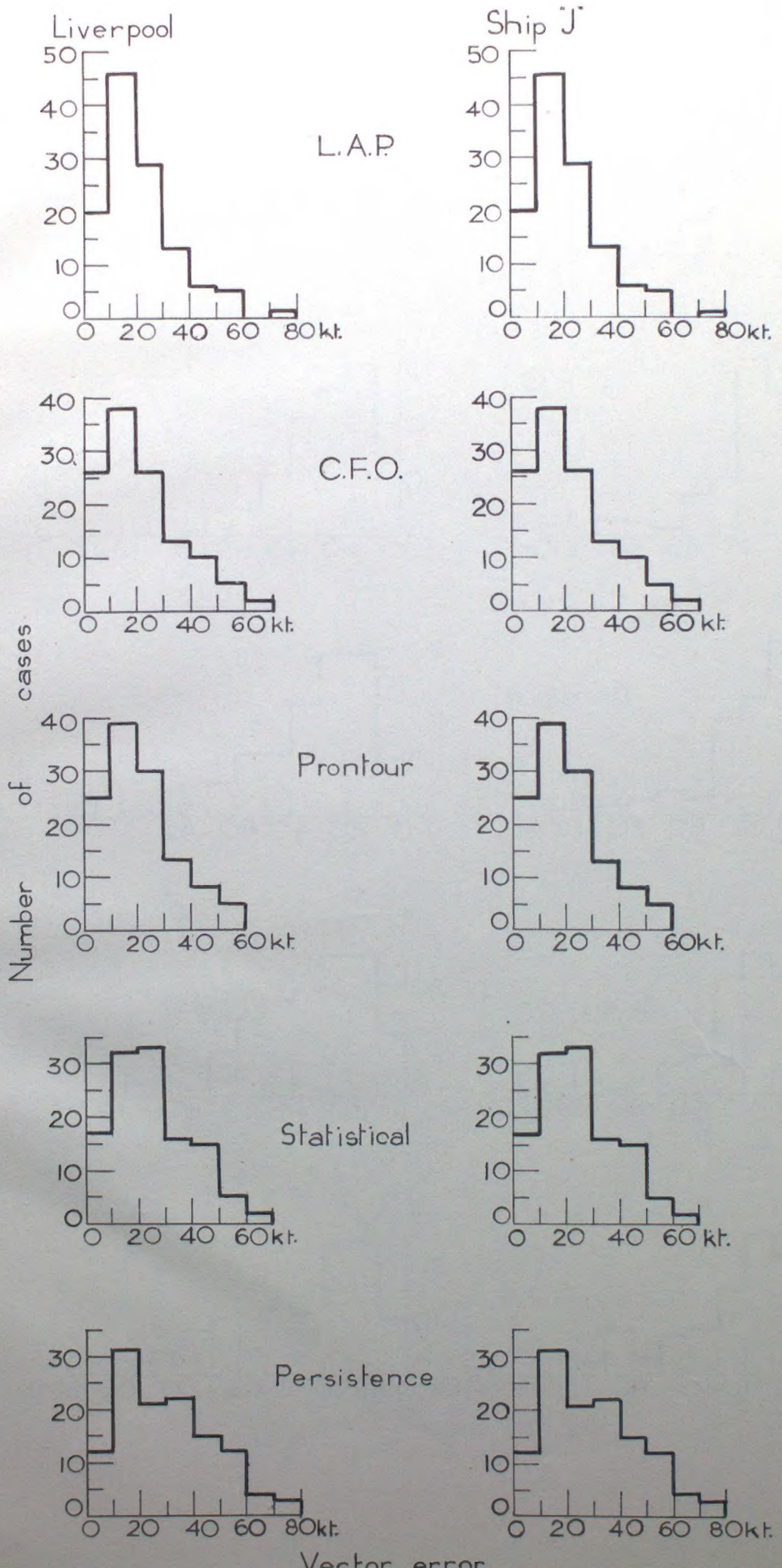


Fig. 4 (c) **M.R.P.953**

Frequencies of forecast errors – point forecasts, 300 mb.
January – April, 1955

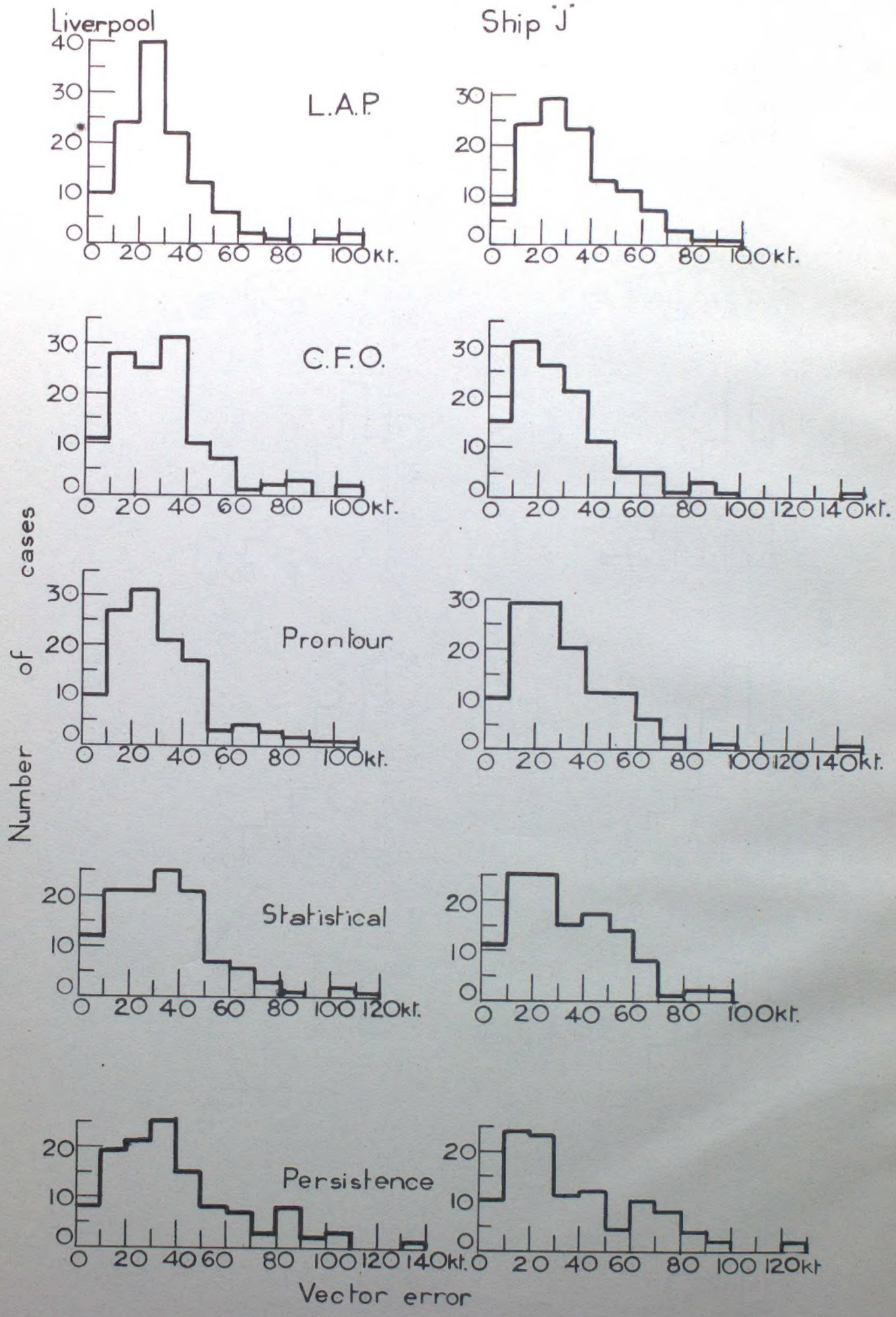


Fig. 4 (d)

Frequencies of forecast errors – point forecasts, 200mb.
January – April, 1955

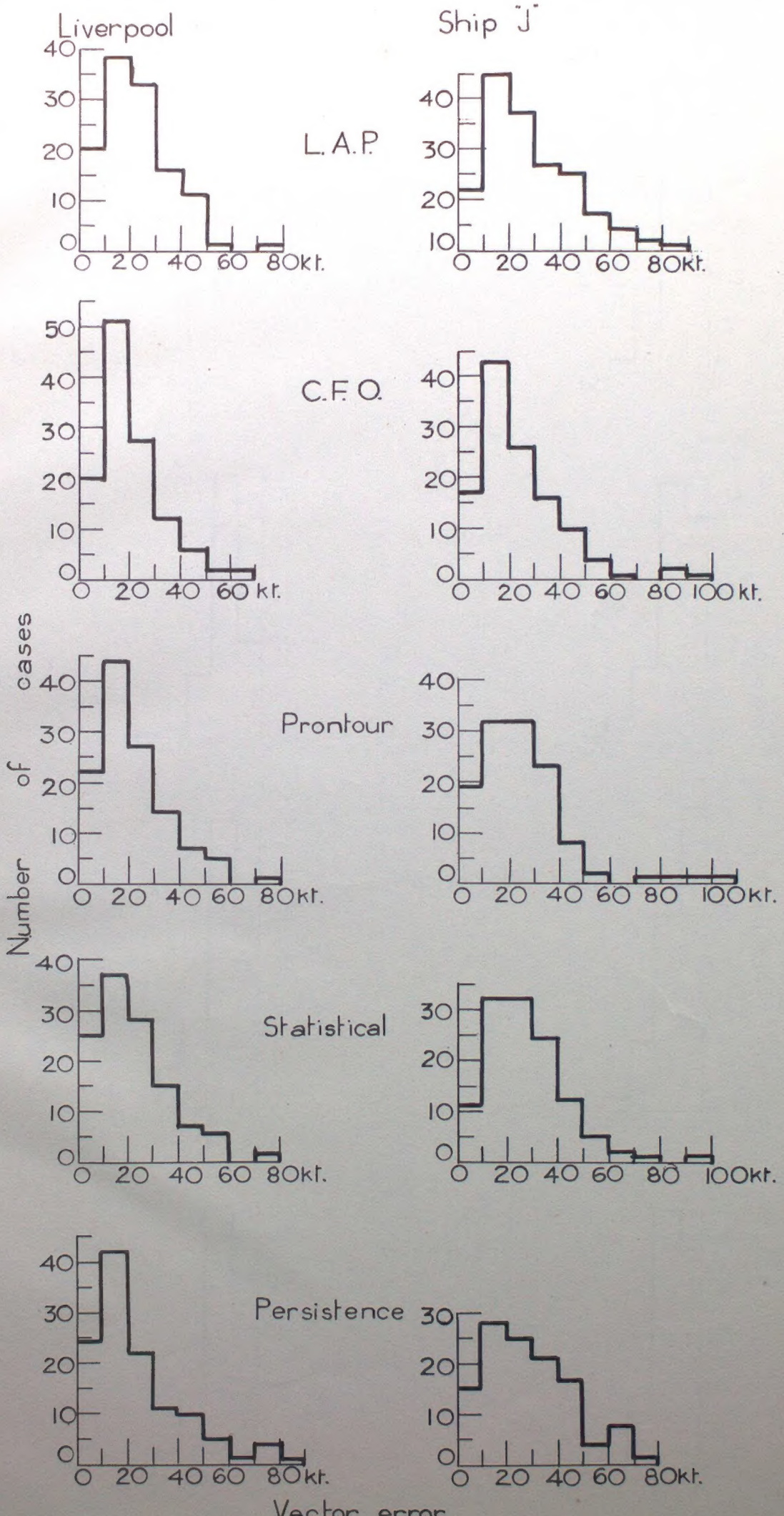
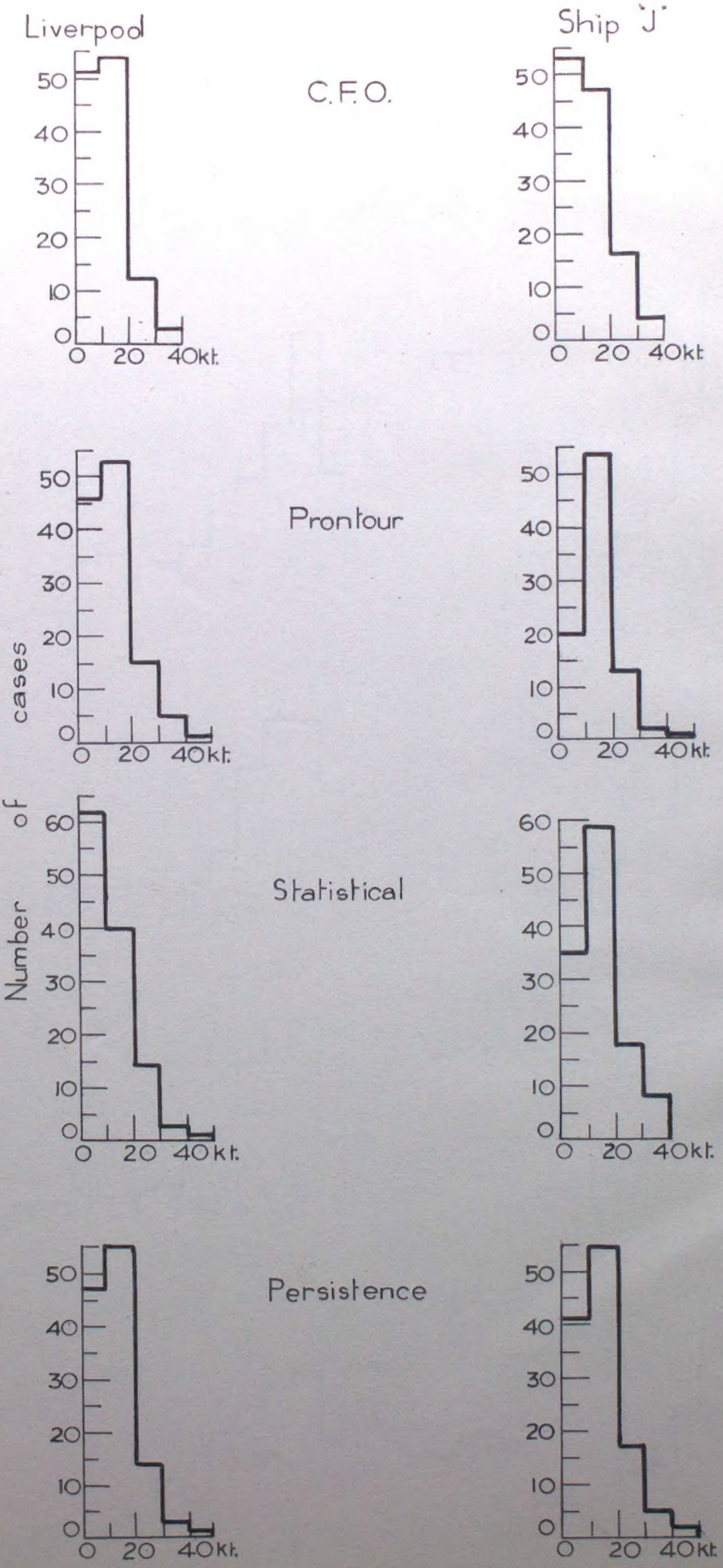


Fig. 4 (e)

Frequencies of forecast errors – point forecasts, 1000mb.
January – April, 1955